MONTHLY NOTICES

OF THE

ROYAL ASTRONOMICAL SOCIETY

Vol. 108 No. 3

MEETING OF 1948 APRIL 9

Professor W. M. H. Greaves, President, in the Chair

The election by the Council of the following Fellows was duly confirmed:

Ida W. Busbridge, M.Sc., M.A., D.Phil., St. Hugh's College, Oxford (proposed by Miss M. G. Adam);

Peter Kirchhoff, Dr. Rer. Nat., 24 Tenth Avenue, Parktown North, Johannesburg (proposed by R. O. Redman);

Anthony James Newberry, 3 Lonsdale Close, Orchard Drive, Edgware, Middlesex (proposed by D. W. Dewhirst);

William Harvey Rayner, 257 Uxbridge Road, Hampton Hill, Middlesex (proposed by W. Rees Wright);

Major Victor Clement Wall, 88 Lavender Grove, Dalston, London, E. 8 (proposed by J. Pickston); and

George Frederick West, Hamal, Newton's Lane, Ruskington, Sleaford, Lincolnshire (proposed by W. L. Shepherd).

The election by the Council of the following Junior Members was duly confirmed:—

Richard Myer Baum, Stocks Lane Post Office, 1 Dee Banks, Boughton, Chester (proposed by H. P. Wilkins); and

John Brian Hutchings, 27 Elmwood Avenue, Hoole, Chester (proposed by H. P. Wilkins).

Sixty-nine presents were announced as having been received since the last meeting, including:—

T. L. Kelly, *Fundamentals of Statistics*, presented by the author; and

H. D. Babcock and C. E. Moore, *The Solar Spectrum*, 6600–13495, presented by the authors.

The President announced that the Council had elected the following as Associates of the Society:—

Dr Walter Baade, of the Mount Wilson Observatory, Pasadena, California, U.S.A.;

Professor Dirk Brouwer, Director of the Yale University Observatory, New Haven, Connecticut, U.S.A.;

Dr Paul W. Merrill, of the Mount Wilson Observatory, Pasadena, California, U.S.A.; and

Dr Bengt Strömgren, Director of the Observatory, Copenhagen.

Professor Bertil Lindblad delivered the George Darwin Lecture (see p. 214), taking as his subject "The Dynamics of the Stellar Systems".

After the lecture, the President presented Professor Lindblad with the Gold Medal of the Society, which had been awarded to him for his work on the rotation of the Galaxy, on the dynamics of stellar systems and on stellar luminosities. The President said:—

"The Gold Medal of our Society, which is the highest honour within our power to bestow, has been awarded to Professor Lindblad for his work on the rotation of the Galaxy, on the dynamics of stellar systems and on stellar luminosities. Our Medallist is both a practical astronomer and a theoretician, and that being so, it is not surprising to find that in testing the dynamical structure which he has erected he has employed the results of his own observational research. In his work on stellar luminosities he has shown that for the later type stars the absolute magnitude is a function of the absorption in the cyanogen bands and of the local colour index near the G band in the spectrum, this colour index being obtained by comparing measures in the spectral regions G to $\lambda 4383$ and $\lambda 4227$ to G. For the earlier spectral types he derived absolute magnitudes from the extended scheme of spectral classification which he himself developed, and so for stars of all types he was able to obtain absolute magnitudes which, when combined with the available apparent magnitudes and proper motions, gave distances and transverse velocities. In this work a low dispersion suffices, so that results can be obtained for a large number of stars, and our Medallist has shown that the data he obtained in this way are consistent with his concept of the Galaxy as an assemblage of rotating sub-systems, the rotation and the dispersion of internal velocities varying with the degree of flattening of the sub-system. By considering the high-velocity stars as forming one of the sub-systems he has explained the tendency of the directions of motions of these stars to avoid the hemisphere whose centre is in the Galactic Plane at longitude 60°. He has also demonstrated that the phenomenon of star streaming is connected with galactic rotation, the direction of preferential stream motion being approximately the direction of the galactic centre.

"In his researches on spiral structure our Medallist began by working out the conditions for the stability of circular orbits near the edge of a highly flattened rotating system of stars. He showed that instability will occur if the flattening is large enough, and that it will develop at any particular point near the edge if the density is small enough, a condition which is likely to be fulfilled if there is a strong radial density gradient in the equatorial plane close to the point in question. If the condition for instability is satisfied, the consequential departure from circular orbits will result in orbits which are initially asymptotic to circular ones. This means that the system will begin to break up by equatorial loss of matter, a result in stellar dynamics which is to some extent analogous to a well-known hydrodynamic result, the equatorial break-up of a rotating fluid mass under certain conditions. Since we believe that the nucleus of a nebular system consists of stars, our Medallist's result provides a dynamical mechanism whereby matter is ejected from the nucleus of a nebula. Perhaps some mathematicians

would have been content to leave it at that, but although our Medallist is a mathematician of high repute he is primarily an astronomer, and he naturally proceeded to consider the form that the ejected matter would assume. In the George Darwin Lecture which he has just delivered he has given us an account of his method of attack and of the results he has obtained, and we must all have been impressed by the success that has been the reward of his efforts to explain the structures that we actually observe.

"I think that anyone who reviews our Medallist's work will get the impression that he attacks problems which are not only astronomically important, but which are intrinsically difficult. As a result of the skill with which he has treated such problems he has been responsible for a substantial advancement in our knowledge. And there is another feature of his career to which it is fitting that I should refer. He has been a source of inspiration to others, and as Director of the Stockholm Observatory he leads a team of which he has good cause to be proud. It would be possible to say a good deal about the fine work which has been done by the members of that team, but we are concerned to-day with its captain.

"Professor Lindblad:—The Gold Medal of the Royal Astronomical Society is not awarded lightly, but only after very careful consideration. This Medal symbolizes the very high respect we have formed of your work in the past, and I may say that it also represents our confidence that the contributions you will make to our Science in the future will enhance that respect. With the Medal our very best wishes go to you."

ON THE DYNAMICS OF STELLAR SYSTEMS

George Darwin Lecture, delivered by Professor Bertil Lindblad on 1948 *April* 9

When talking to the Royal Astronomical Society on the subject of the dynamics of stellar systems, it would be tempting to start with an elaborate sketch of the historical background, beginning in the eighteenth century with the ideas of Thomas Wright of Durham, and with the work of William Herschel. Time does not permit me, however, to indulge in such a wide introduction to my subject, and I must be content to begin with what we may call the modern era of our subject. This era may be said to begin early in the present century with Kapteyn's discovery of the two star streams, and with the subsequent development of the mathematical-statistical methods of describing stellar motions. In a more restricted sense the modern era of our subject begins only thirty years ago with the works of two great pioneers, whom we have all known and admired, and to whom I should like to pay special homage to-day, namely Sir James Jeans and Sir Arthur Eddington. It is true that the contributions of these two giants to the dynamics of stellar systems form only a small part of their total achievements in the science of Astronomy. Nevertheless, their contributions in this field have been of extreme importance to all their successors. Then, new discoveries of capital importance made it possible to gain a deeper insight into the mechanism of our galactic system. The phenomenon of the asymmetrical drift, most thoroughly studied by Strömberg, allowed an interpretation in terms of a general galactic rotation, and Oort's discovery of the differential galactic rotation, which was so soon splendidly confirmed by Plaskett and Pearce, opened up a new field of greatest importance for general researches on the structure of our Galaxy. Since then there have been beautiful contributions to stellar dynamics, especially on the problem of the stochastic effects of stellar passages. Here Chandrasekhar's work on dynamical friction, the time of relaxation and the age of clusters, is of very great interest.

There is, however, another obvious problem of outstanding importance. How can we widen the rotation theory of the Galaxy to a theory of the galaxies in general, which accounts for the characteristic features of the main types of the galaxies? This is a perilous problem, where the risk of making grave mistakes is still considerable, but I have chosen this problem as the main topic of my lecture to-day for several reasons. First, because it is necessary to limit myself to a particular phase of the subject, second, because I have thought that a presentation of the problem, and of a few results which I believe I have obtained, may interest the Royal Astronomical Society, and last, because it is a topic which interests me very much, and which for several years has seldom been entirely out of my thoughts.

Let us consider for a moment Hubble's well-known classification of the nebulae. The series of spiral nebulae begin at the end of the series of elliptic nebulae, and we know by Hubble's observations that the transition has several features of interest, especially as to the appearance of dark matter. We know that there are two parallel series of spirals, the "normal" spirals, and the "barred" spirals, where matter has assembled towards a certain diameter in the equatorial

plane of the system, in such a way that it looks as if the nucleus were threaded on a material "bar". It is clear that a complete theory of the spirals must explain all the main features of this sequence of types.

The spectrographic observations of rotations of nebulae performed at the Lowell, Mount Wilson and Lick observatories have demonstrated that a general motion of rotation is the dominating feature of the motions of the stars in the nebulae of great flattening, just as it is in our Galaxy. There is another feature which we may borrow from the theory of our Galaxy. The phenomenon of the asymmetrical drift may be explained in a general way by a division of our system into sub-systems which have different speeds of rotation, different internal velocity dispersions and different degrees of flattening towards the equatorial plane of symmetry. From the beginning it was evident that these sub-systems were characterized to a considerable extent by different types of population among the stars, i. e. a different mixture of stars with respect to spectral type and absolute luminosity, or by quite special physical characteristics, for instance in the case of planetary nebulae or globular clusters. Considerations of a somewhat similar nature have been advanced recently in a very interesting way by Baade in his segregation of two different types of population, following upon his capital demonstration that the so-called amorphous nebulosity in the spirals may be dissolved into stars. Strongly concentrated towards the equatorial plane are the luminous O, B and A stars, Cepheid variables of long period, open clusters, etc., which will all belong to Baade's population type I. Stars of later spectral types show more complicated conditions. Some of these stars, together with, for instance, the short-period Cepheids, have a distribution of smaller general degree of flattening, but have instead a stronger general concentration to the nuclear regions, and form Baade's population type II, which is also characteristic for the stars in globular clusters. These phenomena, which are of great importance for our ideas concerning stellar evolution, are no doubt connected with the way in which the stars were formed out of an original primitive gaseous medium. In the central layer of a primitive nebula a condensation moving in a nearly circular orbit will move in relative equilibrium with the surrounding matter, which will favour the accumulation of matter into a growing condensation, in contrast to what will be the case farther from the central layer, where a beginning condensation will follow a complicated orbital motion, which carries it almost immediately out of its original surroundings. It seems that we may in this way at least account for a greater relative frequency of very massive stars in the central layer compared with exterior regions. Characteristic differences between stars of one and the same spectral type and the same luminosity group, such as are found to be correlated in our surroundings with the space velocities of the stars, may be due to a systematic variation in the frequency of the elements with the position in the original nebula.

We consider a rotating system of circular symmetry about the centre of gravity in the invariable plane, which we suppose to be a plane of symmetry. The system is supposed to contain gas, small particles and stars, in proportions which need not be specified explicitly beforehand. Small particles and gas will largely occur in the form of individual clouds. If we take out a certain element of volume at a certain moment, the stars in this element will have a certain velocity distribution, dominated by the velocity of rotation. The actual orbits will depend on the nature of the field of force.

We may introduce a coordinate system ξ, η, ζ with origin O following a circular motion in the plane of symmetry with radius r_0 and angular velocity ω_c. The ξ-axis is radial, the η-axis tangential in the direction of rotation, and the ζ-axis at right angles to the plane of symmetry. For small deviations from the circular orbit the motion in ξ, η and the motion in ζ may be considered independent of each other. The motion in ξ, η is of an epicyclic kind (Fig. 1) depending on

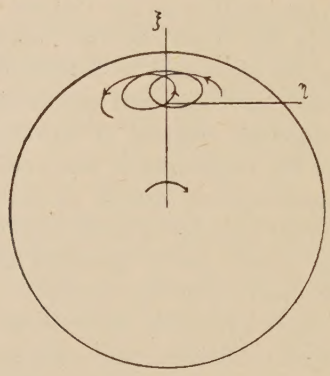


FIG. 1.—*Relative orbit in a rotating coordinate-system following a circular orbit.*

four parameters c, c_0, t_0, t_1, as follows,

$$\left.\begin{aligned} \xi &= \frac{c_0}{2(\omega_c - A_c)} + c\cos\kappa(t-t_0), \\ \eta &= -\frac{A_c}{\omega_c - A_c}c_0(t-t_1) - c\left(\frac{\omega_c}{\omega_c - A_c}\right)^{\frac{1}{2}}\sin\kappa(t-t_0), \end{aligned}\right\} \qquad (1)$$

where

$$\kappa = 2\{\omega_c(\omega_c - A_c)\}^{\frac{1}{2}}.$$

A_c is Oort's constant of differential rotation for the circular motions surrounding $r = r_0$, thus

$$A_c = \frac{1}{4\omega_c}\left(\frac{\partial^2\phi}{\partial r^2} - \frac{1}{r}\frac{\partial\phi}{\partial r}\right),$$

where ϕ is the potential function. The frequency κ of the oscillations in ξ and η is a function of the distance from the centre. It may be assumed to increase inwards and to reach the value $2\omega_c$ in the interior.

As ξ, η are canonical variables, the elements $d\xi\, d\eta\, d\xi'\, d\eta'$ will be constant in the motion, and the general frequency function in ξ, η, ξ', η' may be written

$$f(\xi, \eta, \xi', \eta') = F(c, c_0, t_0, t_1).$$

In a steady state t_0 and t_1 will disappear and a general local distribution giving an ellipsoidal velocity distribution will be

$$F(c, c_0) = C\exp\{-(h^2c^2 + k_1c_0^2 + k_2c_0)\}.$$

We then get the mean motion

$$\xi_0' = 0, \quad \eta_0' = -2\frac{A_ch^2\xi + (\omega_c - A_c)^2(4\omega_ck_1\xi + k_2)}{h^2 + 4k_1(\omega_c - A_c)^2},$$

and the axes α and β of the velocity ellipsoid by

$$\alpha^2 = \frac{\kappa^2}{2h^2}, \quad \beta^2 = \frac{2(\omega_c - A_c)^2}{h^2 + 4k_1(\omega_c - A_c)^2}.$$

These formulae contain the theory of the differential rotation and the velocity ellipsoid for relative velocities which are small compared with the velocity of rotation. In particular, the coefficient A of differential rotation will be

$$A = \omega(1 - \beta^2/\alpha^2).$$

If h^2, k_1, and k_2 were strictly constant within a certain interval of ξ about the origin, the asymmetrical drift S, defined as the difference between the circular velocity and the mean velocity of the group considered, would be

$$S = -\frac{\alpha^2}{2\omega_c}\frac{\partial \log_e \nu}{\partial \xi},$$

where ν is the space density. This agrees with Oort's * formula, when we put $r = \infty$. From more general considerations, starting from the equations of mass motion, and admitting that the effective value of α^2 in the ellipsoidal approximation may change with r, we may derive †

$$S = -\frac{r}{2\Theta_c - S}\alpha^2\left[\frac{\partial \log_e(\nu\alpha^2)}{\partial r} + \frac{1}{r}\left(1 - \frac{\beta^2}{\alpha^2}\right)\right].$$

In certain regions of the system there may be special conditions prevailing. The frequency of oscillation κ may decrease to zero, which means that there will be a kind of asymptotic spiral orbits extending from the region, in which particles may reach great distances from the centre. If we have in a certain region,

$$\kappa^2 < 0,$$

the circular orbits are unstable in this region, between r_1 and r_2, say. To a circular motion with radius $r_1 - \delta_1$ will then correspond a "proper circle" with radius $r_1 + \delta_2$, such that if the particle is moved through the small interval $\delta_1 + \delta_2$ between these two radii, it will from there on depart in an orbit which may extend

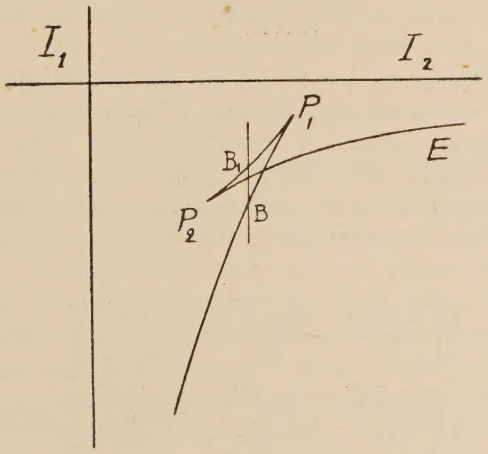


FIG. 2.—*The characteristic diagram of a system.*

widely beyond r_1 and r_2. The conditions for the appearance of a region of instability may be studied by the so-called characteristic diagram, in which the two integrals, the area integral I_2 and the energy integral I_1 are the axes (Fig. 2). A certain envelope curve E in this diagram represents the circular motions and

* J. H. Oort, *B.A.N.* No. 159, p. 281, 1928.

† *Cf.* Astronomical Papers dedicated to Elis Strömgren, p. 123, 1940.

a region of instability corresponds to a singular part of this envelope curve between two singular points P_1 and P_2. The three intersections with the vertical line BB_1 correspond to circular orbits of one and the same I_2. The one corresponding to B_1 is unstable. The conditions are of special interest for spheroids of increasing flattening. Instability begins at the meridian eccentricity $e=0{\cdot}834$, and for infinite flattening the characteristic curve E reaches $I_1=0$, and the asymptotic orbit from the edge extends to infinity.*

The concentration of dark matter into a much flattened formation in the outer regions of the system may contribute effectively to the appearance of a zone of instability. If ϕ is the potential function, and if we define a quantity $\bar{\rho}$ by

$$\partial^2\phi/\partial z^2 = -4\pi G\bar{\rho},$$

we have

$$\kappa^2 = 4\pi G\bar{\rho}\left(\frac{\omega_c^2}{2\pi G\bar{\rho}} + \frac{\rho}{\bar{\rho}} - 1\right),$$

where ρ is the density at the point considered. In order that we shall have $\kappa^2<0$, $(\bar{\rho}/\rho)$ should be comparatively large, which is made possible by a strong concentration towards the equatorial plane combined with a strong radial density-gradient. It seems that the presence of dark matter in the form of small solid particles will favour the appearance of such conditions. In fact, there is an intimate connection between spiral structure and heavy lanes of dark matter, and according to Hubble the first development of spiral structure in the sequence of nebular types is correlated with the occurrence of a peripheral ring of dark matter. In systems of type So where there is no trace of dark matter there is no trace of spiral structure.

Of greatest importance for our problem is a possible instability of the rotating mass as a whole. In order to study this question we use the general equations of mass motion, and start from the conditions of a state of equilibrium. We introduce a coordinate system, x, y, z with the centre of gravity as origin and with the invariable plane, which we suppose to be a plane of symmetry, as x, y-plane, rotating with the angular velocity ω about the z-axis. The density function is $\rho(x, y, z)$ and the gravitation potential $\phi(x, y, z)$. The components of the mean motion at the point x, y, z are u, v, w, and the components of the individual velocities of the particles are U, V, W. We then have the general equations of mass motion

$$\left.\begin{aligned}
\frac{Du}{Dt} - 2\omega v + \frac{1}{\rho}\left[\frac{\partial(\rho\overline{U^2})}{\partial x} + \frac{\partial(\rho\overline{UV})}{\partial y} + \frac{\partial(\rho\overline{UW})}{\partial z}\right] - \omega^2 x &= \frac{\partial\phi}{\partial x},\\
\frac{Dv}{Dt} + 2\omega u + \frac{1}{\rho}\left[\frac{\partial(\rho\overline{UV})}{\partial x} + \frac{\partial(\rho\overline{V^2})}{\partial y} + \frac{\partial(\rho\overline{VW})}{\partial z}\right] - \omega^2 y &= \frac{\partial\phi}{\partial y},\\
\frac{Dw}{Dt} + \frac{1}{\rho}\left[\frac{\partial(\rho\overline{UW})}{\partial x} + \frac{\partial(\rho\overline{VW})}{\partial y} + \frac{\partial(\rho\overline{W^2})}{\partial z}\right] &= \frac{\partial\phi}{\partial z}.
\end{aligned}\right\} \qquad (2)$$

and the equation of continuity

$$\frac{\partial\rho}{\partial t} + \frac{\partial(\rho u)}{\partial x} + \frac{\partial(\rho v)}{\partial y} + \frac{\partial(\rho w)}{\partial z} = 0.$$

* Cf. *Stockholms Observatoriums Annaler*, **15**, No. 4, p. 5, 1948.

In equilibrium we assume the mean motion to be a rotation λr about the z-axis, where r is the distance from this axis, and λ is a function of r and z. We have thus

$$u_0 = -\lambda y, \quad v_0 = \lambda x, \quad w_0 = 0.$$

The components of the dispersion tensor $\overline{U^2}$, $\overline{UV}$, etc., are formed as the weighted means for all particles, whether stars, dust particles, or gas molecules. We assume the dispersion tensor to be characterized by a velocity ellipsoid which has one axis α in the radial direction, one axis β in the direction of the rotation, and the third axis γ in the z-direction. For points on the x-axis we have then

$$\alpha^2 = \overline{U^2}, \quad \beta^2 = \overline{V^2}, \quad \gamma^2 = \overline{W^2}.$$

The assumption that the system advances towards a steady formation may be expressed by the statement that the dispersion tensor approaches such a form. Deviations of the tensor from this form will be associated with certain density variations $\Delta\rho$, which we suppose to be of the first order of small quantities. The direct influence of $\Delta\rho$ on the individual orbits of the particles during a time comparable with the time of revolution of the system will be small, and we may assume that the direct influence of $\Delta\rho$ on the dispersion tensor $\overline{U^2}$, $\overline{UV}$, etc., is of higher order than the first. When considering the further propagation of the density variations $\Delta\rho$ we may therefore assume $\overline{U^2}$, $\overline{UV}$, etc., to be constant and to have the values characterized by the equilibrium form of the system.

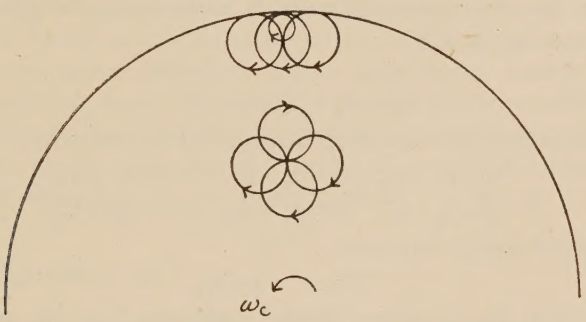


FIG. 3.—*Relative orbits in a quasi-spheroidal system.*

For further illustration we may consider the schematic case of a quasi-spheroidal system, in which we introduce a coordinate system rotating with the angular velocity ω_c of the circular orbits. The relative orbits of free particles, or of individual clouds of gas and small particles, in the plane of symmetry are then circles described with the angular velocity $2\omega_c$ about a local centre. The mean motion and the velocity dispersion at a given point are determined by the amount of matter carried by the orbits intersecting at the point in question. If the dispersion tensor about the mean motion corresponds to the state of a steady formation, local gravitation fields, or local density gradients due to $\Delta\rho$, will only affect the dispersion tensor $\overline{U^2}$, $\overline{UV}$, etc., with terms of higher order. This is immediately clear in the case of the interchange of matter between neighbouring points in the presence of a density gradient. The change in the mean motion caused by the density gradient will not, to the first order, affect the velocity dispersion. In the case of the local gravitation fields we may denote with ϵ_1, ϵ_2, ϵ_3 the disturbances in U, V, W for the particles which are situated in a certain region at a certain moment. If the disturbing field is fairly uniform over regions

comparable with the dimensions of the relative orbits, the residuals ϵ_1, ϵ_2, ϵ_3 will be numerically small compared with the simultaneous effects Δu, Δv, Δw in the mean motion. We may further assume for the accumulated effects of the gravitational disturbances during the orbital motions of the particles that ϵ_1, ϵ_2, ϵ_3 are only weakly correlated with U, V, W, so that in the expressions

$$\overline{(U+\epsilon_1)^2} = \overline{U^2} + \overline{\epsilon_1^2} + 2\overline{U\epsilon_1},$$
$$\overline{(U+\epsilon_1)(V+\epsilon_2)} = \overline{UV} + \overline{\epsilon_1\epsilon_2} + \overline{U\epsilon_2} + \overline{V\epsilon_1}, \text{ etc.},$$

the terms $\overline{U\epsilon_1}$, $\overline{U\epsilon_2}$, $\overline{V\epsilon_1}$, ... may be considered to be of higher order.

In the course of time the small variations of the dispersion tensor may work as viscosity terms in the equations, and transform the kinetic energy of the deviations of the mean motions into an increase of the kinetic energy of the individual motions. As the energy of the deviations of the mean motions

$$\tfrac{1}{2}\int[(\Delta u)^2+(\Delta v)^2+(\Delta w)^2]dm,$$

is of the second order, the resulting effects in $\overline{U^2}$, $\overline{V^2}$, $\overline{W^2}$ will also be of the second order. The derivatives $\partial\overline{U^2}/\partial x$, etc., will thus only vary with small quantities of higher order than the first.

In the study of propagation of the density variations which are caused by an imperfect mixing of the matter of the system we are therefore entitled to assume as an approximation that the dispersion tensor remains constant.

The density function in equilibrium is assumed to be $\rho_0(x, y, z)$ and the gravitation potential $\phi_0(x, y, z)$.

We further define

$$\Omega_0 = \phi_0 + \tfrac{1}{2}\omega^2(x^2+y^2).$$

Considering a variation from the state of equilibrium we write

$$u = u_0 + u_1, \quad v = v_0 + v_1, \quad w = w_0 + w_1$$
$$\rho = \rho_0 + \rho_1, \quad \phi = \phi_0 + \phi_1, \quad \Omega = \Omega_0 + \Omega_1.$$

We have then by Poisson's equation

$$\nabla^2\phi_1 = -4\pi G\rho_1.$$

We may easily form the equations for the variations from the general equations (2). As we are mostly interested in the motions parallel to the equatorial plane, we integrate with respect to the z-coordinate, and form

$$\zeta_0 = \int_{-\infty}^{+\infty}\rho_0\,dz, \qquad \zeta_1 = \int_{-\infty}^{+\infty}\rho_1\,dz,$$

$$\overline{u_1} = \frac{1}{\zeta_0}\int_{-\infty}^{+\infty}u_1\rho_0\,dz, \quad \overline{v_1} = \frac{1}{\zeta_0}\int_{-\infty}^{+\infty}v_1\rho_0\,dz, \quad \overline{\Omega_1} = \frac{1}{\zeta_0}\int_{-\infty}^{+\infty}\Omega_1\rho_0\,dz.$$

We then get the equations

$$\left.\begin{aligned}
&\frac{\partial\overline{u_1}}{\partial t} - \lambda y\frac{\partial\overline{u_1}}{\partial x} + \lambda x\frac{\partial\overline{u_1}}{\partial y} - 2\omega\overline{v_1} = \frac{\partial\overline{\Omega_1}}{\partial x} - \frac{1}{r^2}(\alpha^2x^2+\beta^2y^2)\frac{\partial}{\partial x}\left(\frac{\zeta_1}{\zeta_0}\right)\\
&\qquad\qquad - (\alpha^2-\beta^2)\frac{xy}{r^2}\frac{\partial}{\partial y}\left(\frac{\zeta_1}{\zeta_0}\right).\\
&\frac{\partial\overline{v_1}}{\partial t} - \lambda y\frac{\partial\overline{v_1}}{\partial x} + \lambda x\frac{\partial\overline{v_1}}{\partial y} + 2\omega\overline{u_1} = \frac{\partial\overline{\Omega_1}}{\partial y} - \frac{1}{r^2}(\beta^2x^2+\alpha^2y^2)\frac{\partial}{\partial y}\left(\frac{\zeta_1}{\zeta_0}\right)\\
&\qquad\qquad - (\alpha^2-\beta^2)\frac{xy}{r^2}\frac{\partial}{\partial x}\left(\frac{\zeta_1}{\zeta_0}\right).\\
&\frac{\partial\zeta}{\partial t} + \frac{\partial}{\partial x}(\zeta_0\overline{u_1}) + \frac{\partial}{\partial y}(\zeta_0\overline{v_1}) - y\frac{\partial}{\partial x}(\lambda\zeta_1) + x\frac{\partial}{\partial y}(\lambda\zeta_1) = 0.
\end{aligned}\right\} \qquad (3)$$

We assume that λ is small in the interior of the system, which implies that we have $\alpha^2 = \beta^2$ nearly. We may call this a "quasi-spheroidal" state. We assume further that in the equation of continuity we may consider $\partial\zeta_0/\partial x$ and $\partial\zeta_0/\partial y$ to be small quantities, so that we may ignore $\overline{u_1}\partial\zeta_0/\partial x$ and $\overline{v_1}\partial\zeta_0/\partial y$.

If we define

$$\bar{\psi} = \overline{\Omega_1} - \alpha^2(\zeta_1/\zeta_0),$$

we may write

$$\left.\begin{aligned} \frac{\partial\overline{u_1}}{\partial t} - 2\omega\overline{v_1} &= \frac{\partial\bar{\psi}}{\partial x}, \\ \frac{\partial\overline{v_1}}{\partial t} + 2\omega\overline{u_1} &= \frac{\partial\bar{\psi}}{\partial y}, \\ \frac{\partial\zeta_1}{\partial t} + \zeta_0\left(\frac{\partial\overline{u_1}}{\partial x} + \frac{\partial\overline{v_1}}{\partial y}\right) &= 0. \end{aligned}\right\} \tag{4}$$

These are evidently equations of a type admitting solutions of the character of wave-motions. If we set

$$\eta = \zeta_1/\zeta_0,$$

we get

$$\partial^2\eta/\partial t^2 + 4\omega^2\eta = \alpha^2\nabla_1^2\eta - \nabla_1^2\bar{\phi}_1. \tag{5}$$

As boundary condition we assume that radial motions due to internal density gradients should vanish at the boundary $r = a$. This is in a way a schematic assumption, as the actual boundary region must cover a certain interval in r with a finite density gradient. If there is a density gradient sufficient to produce an exterior region of instability, in the way described above, the region of this density gradient may be rather narrow and its effect on the interior motions may be appropriately defined in the way just described. If we set

$$\eta \sim e^{i\sigma t},$$

we should then have for $r = a$

$$r\frac{\partial\eta}{\partial r} - \frac{2i\omega}{\sigma}\frac{\partial\eta}{\partial\theta} = 0. \tag{6}$$

In the quasi-spheroidal case deformations of the system at constant density should give approximately $\nabla_1^2\bar{\phi}_1 = 0$. For the potential due to density variations we have

$$\nabla^2\phi_1 = -4\pi G\rho_1,$$

and introduce a function f by setting

$$\nabla_1^2\phi_1 = -4\pi Gf\rho_1.$$

Coutrez * has shown that the function f has the important property

$$\nabla_1^2\rho_1 = \nabla^2(f\rho_1).$$

We assume now that ρ_1 and $\nabla_1^2\rho_1$ are of opposite sign, and that, if we write

$$\frac{\partial^2\rho_1}{\partial z^2} = \frac{1-\mu}{\mu}\nabla_1^2\rho_1,$$

μ will be approximately constant, with $0 < \mu < 1$, from the equatorial plane $z = 0$, where $\partial\rho_1/\partial z = 0$, up to a certain closed surface $z = g(x, y)$, on which $\rho_1 = 0$. We

* Coutrez, *Stockholms Observatoriums Annaler*, **15**, No. 3, 1947.

assume that ρ_1 vanishes outside this surface. Inside the surface we have then

$$\nabla^2\rho_1 = \nabla_1^2\rho_1 + \frac{\partial^2\rho_1}{\partial z^2} = \frac{1}{\mu}\nabla_1^2\rho_1,$$

and thus

$$\nabla^2(\mu\rho_1 - f\rho_1) = 0.$$

As ρ_1 vanishes on the surface $z = g(x, y)$, we must have inside this surface

$$f = \mu.$$

If ρ_1/ρ_0 is fairly independent of z, so that we may put $\rho_1/\rho_0 = \zeta_1/\zeta_0$, we may write

$$\nabla_1^2\overline{\phi_1} = -4\pi G\overline{f\rho_0}\eta,$$

and thus

$$\partial^2\eta/\partial t^2 + 4\omega_1^2\eta = \alpha^2\nabla_1^2\eta,$$

where

$$\omega_1^2 = \omega^2 - \pi G\overline{f\rho_0}.$$

For $\eta \sim e^{i\sigma t}$ we have then

$$\nabla_1^2\eta + k^2\eta = 0; \quad \alpha^2k^2 = \sigma^2 - 4\omega_1^2.$$

Introducing polar coordinates r, θ and assuming that the angular variable enters in η by the factor $e^{is\theta}$, where s is an integral number, we have

$$\frac{\partial^2\eta}{\partial r^2} + \frac{1}{r}\frac{\partial\eta}{\partial r} + \left(k^2 - \frac{s^2}{r^2}\right)\eta = 0.$$

The solution which is finite for $r = 0$ is

$$\eta = AJ_s(kr)e^{i(\sigma t + s\theta)},$$

with the boundary condition

$$kaJ_s'(ka) + (2s\omega/\sigma)J_s(ka) = 0.$$

The conditions are formally similar to those of a rotating sheet of water within a circular boundary.* We may set

$$x = k^2a^2, \quad y = \sigma/2\omega, \quad \beta = 4\omega^2a^2/\alpha^2,$$

and then have

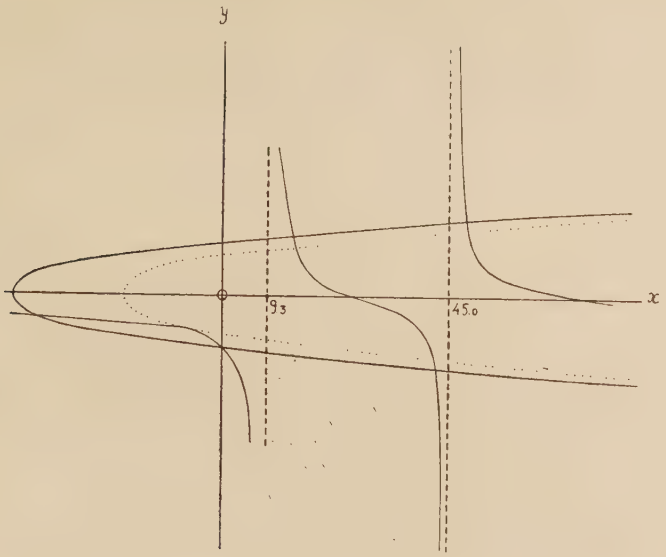
$$\left.\begin{aligned} y^2 &= (\omega_1/\omega)^2 + x/\beta, \\ y &= -\phi(x) = -sJ_s(ka)/kaJ_s'(ka). \end{aligned}\right\}$$

For a given s the function $y = -\phi(x)$ has several branches, and the solutions are obtained by the intersections of these branches with the parabola which represents the first equation. In Fig. 4 the full-drawn parabola is valid for $\omega_1 = \omega$, whereas the dotted parabola corresponds to the case $\omega_1 < \omega$. The scale of y is ten times that of x.

Most of the waves will have $x > 0$ and a frequency close to $\sigma = \pm 2\omega$. Under the tidal forces of a developing arm structure, which will have approximately the frequency 2ω, forced waves of these types will play a part in the actual segregation of a system into a normal type of spiral.

Waves of special interest are those corresponding to intersections between the parabola and the branch of $y = -\phi(x)$ which extends towards negative values

* *Cf.* Lamb, *Hydrodynamics* (Cambridge, 1924), p. 301.



5 · 2

FIG. 4.

of x. In this case k^2 is negative, and the amplitude function of the wave is positive throughout. If the vertex of the parabola lies fairly close to the origin, the equations may give complex values of x and y. A complex value of σ will mean that the wave is unstable, with an amplitude varying exponentially with the time. We have for the vertex of the parabola

$$x_0 = -\beta(\omega_1/\omega)^2 = -4\omega_1^2 a^2/\alpha^2.$$

If the general flattening towards the equatorial plane is great, so that the numerical value of $\partial^2\rho_1/\partial z^2$ is great compared with that of $\nabla_1^2\rho_1$, the value of f will be small and ω_1 will be close to ω. In this case the velocity dispersion α must be small and the numerical value of x_0 will be great. With increasing f the value of ω_1 will decrease and α must increase at the same time. It seems, therefore, that the instability of the wave may take place if f and α are not too small quantities. We shall return presently to this case in connection with the discussion of the barred spiral nebulae.

Variations of density, with a minimum of simultaneous deformations, constitute a limiting case of possible wave-motions. Other limiting cases of wave-motions are a wave of general expansion and contraction and waves of pure deformation of the boundary, without internal density gradients. Together with the density waves these limiting cases may be represented symbolically by the sides of a triangle enclosing possible types of variation. A point within the triangle may be assumed to represent a certain harmonic motion, and the inverse ratios of the distances from the three sides may indicate its general character. The actual state of motion will be a superposition of various harmonic motions. In the quasi-spheroidal case, where the deformations give nearly $\nabla_1^2\overline{\phi}_1 = 0$, the coupling between the three types will on the whole not be very strong, and the motion may dissolve essentially into the harmonic motions of the three limiting types.

The motion of expansion and contraction for a quasi-spheroidal system of great flattening may be assumed to be slow. There remains then the very

important waves of deformation. If $\zeta_1 = 0$ we obtain readily

$$\nabla_1^2 \Omega_1 = 0,$$

and the solution within a boundary circle $r = a$ may be written

$$\Omega_1 = A_n r^n e^{i(\sigma t + n\theta)}.$$

The radial displacement will be

$$\delta r = \alpha_n r^{n-1} e^{i(\sigma t + n\theta)},$$

where

$$\alpha_n = \frac{nA_n}{\sigma(2\omega - \sigma)}.$$

The first non-trivial case will be $n = 2$. In order to derive Ω_1 from the radial displacements δr we assume that for this purpose the system may be represented by a homogeneous spheroid of a certain degree of flattening determined by the eccentricity e of the meridian section. We then derive readily * for $n = 2$

$$A_2 = \alpha_2 \omega^2 F(e),$$

where $F(e)$ is a certain function of the eccentricity e which may be found from the elementary formulae for the potential of homogeneous ellipsoids. This gives

$$\sigma = \omega\{1 \pm (1 - 2F(e))^{\frac{1}{2}}\}.$$

With increasing e the frequency σ passes from real to complex values at an eccentricity e_1 determined by

$$F(e_1) = \tfrac{1}{2}.$$

Introducing $\zeta = (e^{-2} - 1)^{\frac{1}{2}}$, we find

$$(3\zeta_1^4 + 8\zeta_1^2 + 1) \operatorname{arc\,cot} \zeta_1 - 3\zeta_1^3 - 7\zeta_1 = 0.$$

The wave in question is the one which was found in 1860 by Riemann for a rotating Maclaurin spheroid. It consists in a prolonging of the system along a certain diameter and a shortening along the diameter at right angles to this one, the polar axis of rotation remaining unchanged. The equation in ζ_1 agrees with the results of Riemann † and Bryan.‡ The instability takes place when the eccentricity of the meridian section exceeds 0·9529, i. e. when the ratio of the axes falls below 0·30. The real part of σ is equal to ω, and the wave may be written

$$\Omega_1 = Ce^{\gamma t} r^2 \cos(\omega t + 2\theta),$$

which proceeds in θ with half the angular speed of the system itself.

The field of force which carries the wave has not, however, the same effect on all particles of the system. Introducing the rotating coordinate system ξ, η following the circular motion ω_c at $r = r_0$, and limiting ourselves to the effects of the first order, we have

$$\left.\begin{aligned} \frac{d^2\xi}{dt^2} - 2\omega_c \frac{d\eta}{dt} - 4\omega_c A_c \xi &= \frac{\partial \phi_1}{\partial \xi}, \\ \frac{d^2\eta}{dt^2} + 2\omega_c \frac{d\xi}{dt} &= \frac{\partial \phi_1}{\partial \eta}. \end{aligned}\right\} \qquad (7)$$

* Cf. *Stockholms Observatoriums Annaler*, **15**, No. 4, p. 20, 1947.

† B. Riemann, *Abhandlungen der K. Ges. der Wiss. zu Göttingen*, Math. Klasse, IX, 3, 1860.

‡ G. H. Bryan, *Phil. Trans. Roy. Soc.* A, **180**, 187, 1888.

Polar coordinates relative to the centre of gravity in a coordinate system rotating with the angular speed ω_c may be denoted by r, θ'. Comparing with the coordinate system of angular speed ω, in which the polar coordinates are r, θ, we have

$$\omega t + 2\theta = \omega' t + 2\theta',$$

where

$$\omega' = 2\omega_c - \omega.$$

We shall then have

$$\phi_1 = C_0 e^{\gamma t} r^2 \cos(\omega' t + 2\theta').$$

If γ is small, we get the solution

$$\left.\begin{aligned} \xi &= \frac{4\omega_c + 2\omega'}{\omega'(\kappa^2 - \omega'^2)} C_0 r_0 \cos(\omega' t + 2\theta'), \\ \eta &= -\frac{2\omega'^2 + 4\omega'\omega_c + 8\omega_c A_c}{\omega'^2(\kappa^2 - \omega'^2)} C_0 r_0 \sin(\omega' t + 2\theta'), \end{aligned}\right\} \quad (8)$$

where $\kappa^2 = 2\omega_c(\omega_c - A_c)$. For $\kappa = \omega'$ there is resonance with the free radial oscillations, and the effects in ξ and η are large. κ may be assumed to vary from the value $2\omega_c$ in the interior to zero at the edge. The resonance $\kappa = \omega'$ will occur fairly close to the edge, for $r = r_1$, say. For κ slightly larger than ω', the

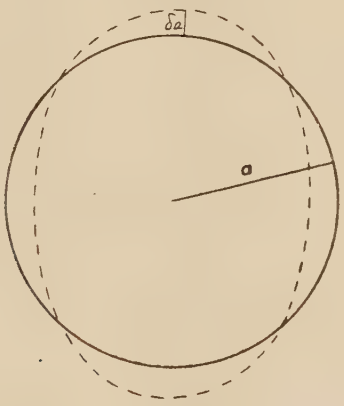


FIG. 5.—*Deformation of the system by the Riemann wave.*

tide is direct and for κ slightly smaller than ω' the tide is indirect. Particles in relative orbits which have their centres inside the circle $r = r_1$ will thus be outgoing at the same time as particles in relative orbits with centres outside the circle $r = r_1$ will be in-going. It seems, therefore, quite possible that the former particles of large amplitude may penetrate the outer parts of the boundary region and attain their "proper circles", from which they will depart in the asymptotic orbits to wide distances from the system.

Because the ejection follows the Riemann wave, with half the angular speed of the system, the locus of the ejected particles, however, will not coincide with the asymptotic orbit in a fixed coordinate system. We assume schematically that a particle moving with the angular speed ω_0 at the edge $r = a$ departs in an asymptotic orbit. We then have for the orbital motion in a fixed coordinate system

$$d\theta/dt = \omega_0 a^2/r^2.$$

We now introduce a coordinate system moving with the angular speed $\frac{1}{2}\omega$. For the present purpose we may assume $\omega_0 = \omega$. If r, θ'' are the coordinates of the particles in this system, we have

$$\frac{d\theta''}{dt} = \omega_0\left(\frac{a^2}{r^2} - \frac{1}{2}\right),$$

and

$$\frac{d\theta''}{d\theta} = 1 - \frac{r^2}{2a^2}.$$

We then derive

$$\theta'' = -\frac{1}{2a^2}\int_{\theta_1}^{\theta}(r^2 - 2a^2)d\theta,$$

where $\theta = \theta_1$ marks the turning-point of the orbit for which

$$r = a\sqrt{2}.$$

For a given asymptotic orbit the curve r, θ'' may readily be computed by numerical integration. It is evident that the locus of the ejected particles must follow the curves r, θ'' in the coordinate system moving with the angular speed $\frac{1}{2}\omega$. The locus of the ejected particles will thus have a turning-point at the distance $r = a\sqrt{2}$ from the centre. From this point on, the locus will lag behind in the rotation. For want of a better term we shall call the two opposite loci of ejected matter "singular" spiral arms.

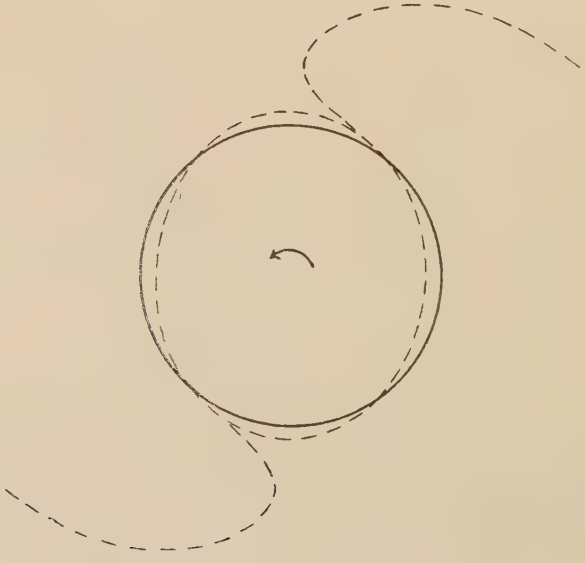


FIG. 6.—*The singular arms formed by matter ejected by the Riemann wave.*

This ejection of matter means a marked deviation from circular symmetry in the general distribution of mass in the system. The matter situated at the turning-point of a singular arm moves with a speed of rotation equal to that of the free wave, and the disturbing effect on the rotating system may be shown to be very strong. The disturbing force of the ejected matter will therefore start a new wave along the edge of the system. The energy integral

$I_1 = u^2 + v^2 + \omega^2 - 2\phi$, where u, v, w denote for the moment the velocity components in a coordinate system fixed in space, will obey the relation

$$dI_1/dt = -2\partial\phi/\partial t.$$

Along the edge immediately after a singular arm in the rotation we shall have $\partial\phi/\partial t < 0$ and immediately before the arm $\partial\phi/\partial t > 0$. The disturbance therefore acts in such a way that the energy of the particles which follow immediately after the singular arms in the rotation will increase, whereas the energy of particles which immediately precede the arms in the rotation will decrease. This means that the disturbance will lift the matter which follows behind the singular arms in the rotation, and press down the matter which precedes the singular arms.

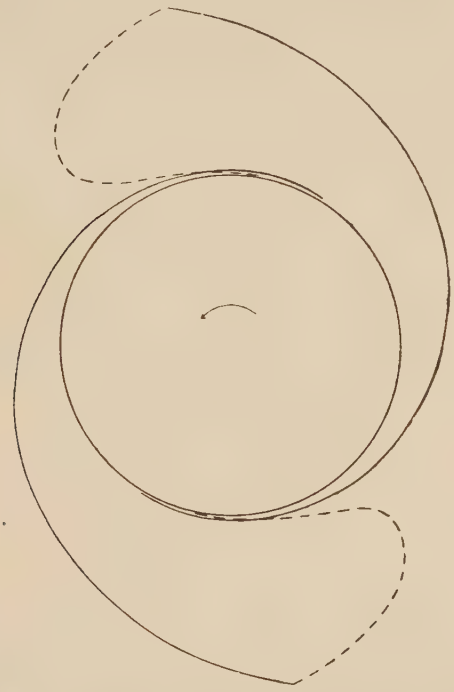


FIG. 7.—*Typical relation between singular arms (dotted curves) and normal arms (full-drawn curves).*

In a coordinate system following the mean angular velocity ω of the system, this wave, which we may call the "dissolving wave", will be propagated along the edge against the rotation with a higher angular speed than that of the turning points of the singular arms. It will proceed by the field of disturbance of the entire mass of displaced matter, and must therefore at first be steadily accelerated in the direction against the rotation. We may assume that this acceleration takes place until the disturbance will be in resonance with the radial oscillations of the particles in the interior of the system. When this state is attained, the points of segregation will proceed against the rotation with the angular speed ω, so that the points of segregation will then be actually fixed in space. After a displacement of 180° relative to the "singular" arm, the dissolving wave will receive a new impulse by the opposite singular arm, and so on. The front of disturbed matter which proceeds against the rotation is assumed to represent the "normal" arms of a spiral nebula. The singular arms will appear more or less as envelope curves of the end-points of various branches of the normal arms.

A most important comparison between theory and observation is afforded by the spiral nebula M 81. At the side of the well-defined normal spiral arms it is possible to trace here in two diametrically opposite regions a diffuse structure, made up of a multitude of curved arcs of matter, which corresponds to the "singular" arms ejected by the Riemann wave. In the theoretical picture the normal spiral arms are matched against asymptotic orbits extending from a quasi-spheroidal mass in which the small axis is equal to 0·26 of the equatorial radius. It is further assumed that the equatorial plane of the nebula forms an angle of 30° with the line of sight. The relation between the "singular" and "normal" arms is further determined by the circumstance that these curves should intersect for a moving particle near the ends of the normal arms. It is actually found that the theoretical singular arms correspond in the mean very well to the structure identified as representing the singular arm structure in the nebula.

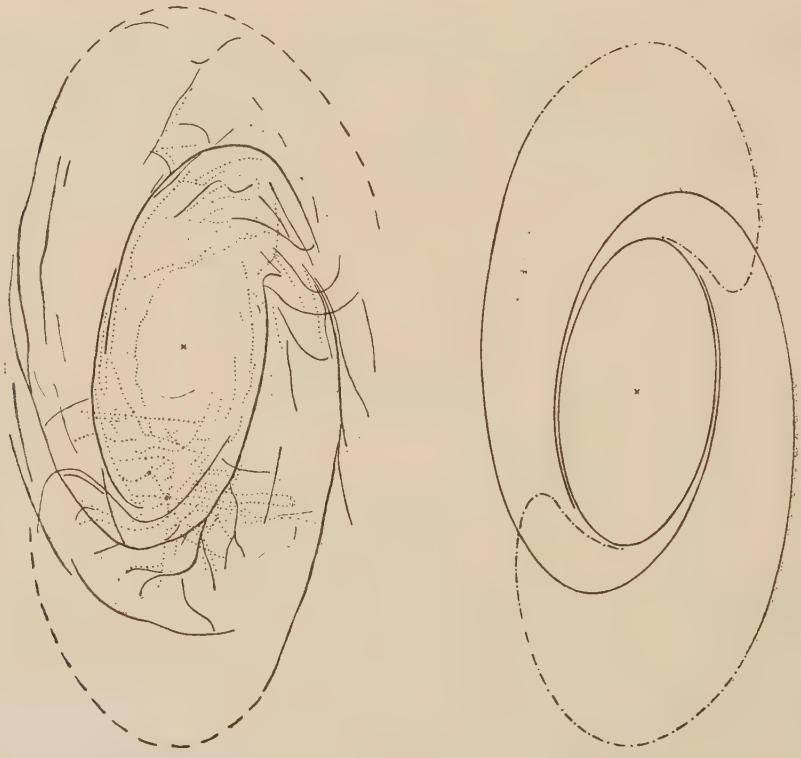


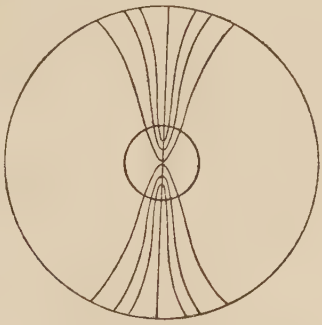
FIG. 8.—*Observed and theoretical spiral structure in M* 81.

The agreement between the normal arms and the asymptotic orbits in M 81 is due to the circumstance that the dissolving wave disintegrates here a ring of matter situated close to the original edge of the system. When the dissolving wave approaches towards the centre of the nebula, the orbits of the deflected particles will no longer be of the asymptotic kind, but rather parts of ellipses with centres coinciding with the centre of the system. A rough calculation shows that we may actually expect in this case a formation of arms which approach to logarithmic spiral curves of the kind actually observed in nebulae of well-developed spiral structure. Moreover, the disintegration will be most effective

and conspicuous for the "sub-system" of strong flattening and of maximum angular speed of rotation roughly corresponding to Baade's population type I. On account of the small velocity dispersion in this sub-system, the particles will crowd together in the arm structure and will be less easily dispersed than particles of more widespread sub-systems of larger velocity dispersion.

The age of the present spiral structure of M 81 may easily be expressed in the period of revolution as unit of time. The angular interval between the points of ejection in corresponding normal and singular arms is about 300°. As the angular difference between these points increases with half the angular speed of the system, it follows that the present spiral structure should have developed during somewhat less than two whole revolutions of the system.

A nebula in which the singular spiral arms are very well developed is NGC 2681. Here, this structure dominates the outer, faint parts of the nebula, whereas the dissolving wave has penetrated to an inner level in the system. The inner, normal spiral arms are therefore surrounded by a diffuse outer spiral structure, in which the direction of winding is opposite to that of the inner arms. The theoretical run of the singular arms has been computed for a ratio of the axes of the effective central system equal to 0·20.



$\Delta\rho \sim \rho I_2(\kappa r) \cos(\sigma t + 2\theta)$

FIG. 9.—*The density waves* $s=2$, $k^2<0$, *which forms the "bar".*

The interrelation between singular and normal spiral arms may have important consequences for the interpretation of certain phenomena in our own Galaxy. When the singular arms dissolve gradually, the corresponding particles will follow orbits of wide amplitude in the radius vector. The matter of these arms, which comes to a considerable extent from regions of fairly large disturbances from the equatorial plane, may be expected to contain largely matter of Baade's population type II, whereas the normal arms will contain preferably stars of population type I. We may therefore expect that the outer parts of the spiral structure will be traversed by high-velocity stars partly belonging to population type II. This gives a new point of view concerning the interpretation of the high-velocity stars in our neighbourhood of the galactic system.

We now pass on for the moment to a discussion of the series of barred spirals in Hubble's classification. The original disturbing wave in this case cannot very well be the Riemann wave, but appears to be a density wave in which matter assembles towards a certain diameter in the system, and disperses about the diameter at right angles to this one. We shall assume that this density wave may be identified with an unstable density wave $s=2$, $k^2<0$, which has been discussed above in relation to Fig. 4. The real part of σ has a small negative value, and

accordingly the wave has an angular speed of rotation which exceeds slightly that of the system. We may therefore assume that the wave follows fairly closely the angular speed of the circular orbits.

We have seen that this wave is likely to be unstable if f and α are not too small quantities, which means that the general flattening of the system is not too great. An important point is that we may account in this way for the fact that the barred spirals have a maximum frequency at an earlier sub-division of the spiral type than the normal spirals.

The density wave in question explains directly the formation of the "bar" along the diameter of maximum frequency. The disturbance exerted by this kind of formation on the internal motions of particles is of great interest. In the neighbourhood of the bar, we may assume that the disturbing force is directed towards the bar, but that it is otherwise uniform. We assume further that the force increases proportionally to e^{kt}. Introducing a rotating coordinate system ξ, η with the origin O following the angular speed ω_c at $r=r_0$, and with the radial ξ-axis forming an angle v with the bar, we have the components of force $-Be^{kt}\sin v$ and $Be^{kt}\cos v$, where B is a constant. For small ξ, η we have then the equations of motion, assuming $A_c=0$,

$$\left.\begin{aligned} \frac{d^2\xi}{dt^2}-2\omega_c\frac{d\eta}{dt} &= -Be^{kt}\sin v, \\ \frac{d^2\eta}{dt^2}+2\omega_c\frac{d\xi}{dt} &= Be^{kt}\cos v. \end{aligned}\right\} \tag{9}$$

The solution starting from $\xi=0$, $\eta=0$ for $t=-\infty$ then is

$$\left.\begin{aligned} \xi &= \frac{2\omega_c\cos v-k\sin v}{k(k^2+4\omega_c^2)}Be^{kt}, \\ \eta &= \frac{2\omega_c\sin v+k\cos v}{k(k^2+4\omega_c^2)}Be^{kt}. \end{aligned}\right\} \tag{10}$$

We have then

$$\frac{\xi}{\eta}=\frac{2\omega_c\cos v-k\sin v}{2\omega_c\sin v+k\cos v}.$$

If k is small the effects in ξ and η at the present time, which we may assume to be $t=0$, are appreciable, even if B is small. In this case we have, moreover,

$$\xi/\eta=\cot v.$$

This shows that the displacement will be along the bar, outwards in the quadrants following the bar in the rotation and inwards in the preceding quadrants. If the bar is vertical, and the rotation occurs in the positive, anti-clockwise direction, the right-hand part of the nebula close to the bar will move upwards and the left-hand part downwards.

That the motion will be of this general kind may also be seen in the way used above for the normal spirals. It is evident that the energy integral I_1 will increase for particles immediately following the bar, and decrease for particles immediately preceding the bar. We must expect that an ejection of matter from the edge of the system begins in the region immediately following the bar. The detached matter lags behind in the rotation and propagates a disturbance along the edge,

but with fairly poor effect, until the disturbance reaches again the regions immediately following the bar. It will then combine in a characteristic way with the disturbing effect of the bar which has just been described.

The particular way in which the bar disturbs the motion in its surroundings explains in a very direct way several characteristic features of the barred spirals. We may assume that systems of this kind have originally consisted of a vast central system of rotational symmetry surrounded by one or several annular formations in the equatorial plane. A slow growth of the bar will cause deflections of the kind described above, and the outer rings will be broken by these deflections in the characteristic way which is particularly well illustrated in the beautiful nebula NGC 1300, reproduced on Plate 5. Even much more complicated structures presented by nebulae like NGC 1398 may be explained in a general way.*

A third type of spiral system is presented by the case when the system is a component in a group of two or more systems, and when there are plain indications that there is a considerable disturbance of the system by its neighbours. It appears in some such cases, e.g. NGC 3718 (disturbed by NGC 3729) and NGC 3190 (disturbed by NGC 3187), that there occurs an enormous development of the singular arm structure, thanks to the disturbance by the component. This disturbance will at times coincide in phase with the disturbance by the Riemann wave. A coincidence of the out-going nodes of the two kinds of wave-motions will occur along a certain diameter in the rotating system, and may cause an unusually strong ejection of matter at the ends of this diameter at the times of coincidence. In NGC 3718 this diameter is likely to be the one along which there appears a pair of strong radial lanes of dark matter.

Very rich spiral nebulae of late type like M 101 show a great many details which are of extreme interest for the dynamical theory. Even here the run of the singular arms may be traced, partly by the run of certain dark lanes, partly as envelope curves of the end-points of the various branches of the ordinary spiral arms. On the whole the spiral structure of M 101 is clearly of a fan-shaped type and it appears that all the detached masses of which the structure consists have originated from a compact mass by relative displacements which are in general agreement with the theory.

The same conclusions are applicable to the nebula M 33, where the arguments may be followed up to some extent by using the radial velocities of gaseous nebulae in the spiral determined by Mayall and Aller. This material, together with that derived by H. W. Babcock for gaseous nebulae in M 31, constitutes material of greatest importance for theoretical studies of the internal motions in spiral nebulae. In M 33 Mayall and Aller derived a general motion of rotation, and Oort has shown that there is probably a general motion of expansion in the main body of the spiral. Objects situated in successive branches of one and the same spiral arm give some information which appears to confirm in a general way the theoretical demand that the rotation occurs in the direction of winding of the spiral arms when these are followed from the centre outwards. From the observed velocities and the present angular distances between the branches we conclude that the age of the present spiral structure is comparable with the time of revolution of the nebula, of the order of 10^8 years.

The present theory has certain immediate consequences as to the general structure of the spiral systems, for instance concerning the relative distribution

* Cf. *Publ. A.S.P.*, **59**, 305, 1947.

of bright and dark matter. The pictures of NGC 4565 and NGC 7331 in Plate 4 illustrate a certain apparent paradox concerning the appearance of nebulae of great inclination against the line of sight, a paradox which seems to be responsible for a great deal of the difference of opinion regarding the direction of rotation in the nebulae.

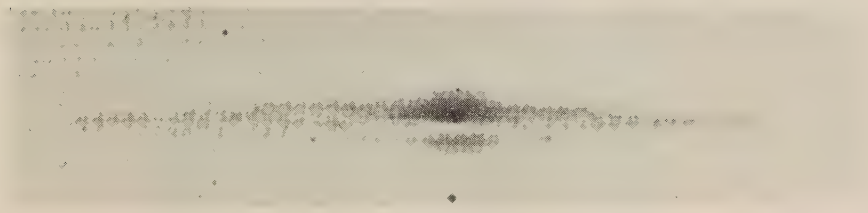
In the case of the upper nebula, NGC 4565, we know with certainty that the lower part of the nebula is the near one and that the right-hand part is approaching in the rotation (observation by Mayall kindly communicated to the author). Our theory predicts that the windings of the arms should open up in the clockwise direction, as is obviously the case with the lower nebula, NGC 7331. This conclusion as to the direction of winding of the spiral arms of NGC 4565 is rather well confirmed by a direct and detailed comparison between the spiral structures of the two nebulae; it is of special importance in this connection to consider the contour of the nebular disk of NGC 4565 in relation to the "nodal line" of the nebula, i.e. the intersection between the nebular plane and the celestial sphere, which may be drawn with fair accuracy from the positions of the vertices of the inner arms. In this way the conditions of NGC 4565 support the theoretical conclusions.

For NGC 7331 we know from spectrographic observations that the right-hand part is approaching in the rotation. We conclude that in this case there is agreement with our theory, if the lower part of the nebula is the near one.

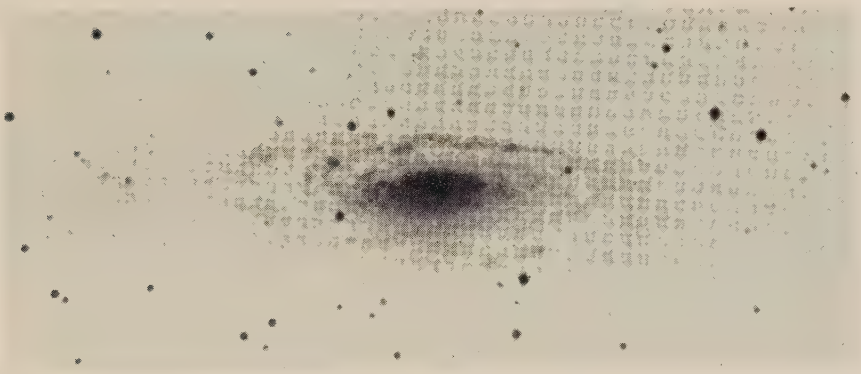
When we compare the two nebulae superficially, especially if we use reproductions on which the inner central parts are over-exposed so that the details are largely obliterated, there will seem to be a radical difference as to the general appearance of the central regions. NGC 4565 has a well-marked central bulge, which seems to be absent, or rather turned the opposite way, in NGC 7331. The apparent difference between the two nebulae, however, will be largely explained when we consider the two pictures of the central regions of NGC 7331. The picture to the left has been taken in ultra-violet light, the picture to the right in red light. It is evident that the central bulge is actually very much smaller in NGC 7331 than in NGC 4565, and that the spiral arms approach very close to the centre in the former nebula. The photometry in ultra-violet and red has allowed us to study the isophotes and isochromes in the central region.* Our conclusion is that the central region of NGC 7331 contains a considerable amount of dark matter which obliterates a great deal of the spiral structure on the side which is turned away from the observer. It is further probable that the light from the intense central region is to a considerable extent subject to diffusion in this dark matter. As the diffusion is likely to be strongly "forward-throwing", i.e. to favour small angles between the incident and out-going beams, it will tend to make the near side apparently more luminous than the farther side. Accurate investigations concerning the polarization of light about the centre, similar to those carried out by Öhman for the central region of M 31, are planned for this nebula, at the Stockholm Observatory, and may be expected to aid in determining the amount of diffusion.

The peculiar character of the absorption of light in the central regions of the late-type spirals may be illustrated by two pictures of the nebula NGC 5055, one taken in red light and the other in blue light (Plate 5). There is here a kind of G-shaped dark lane which actually marks a discontinuity in the distribution

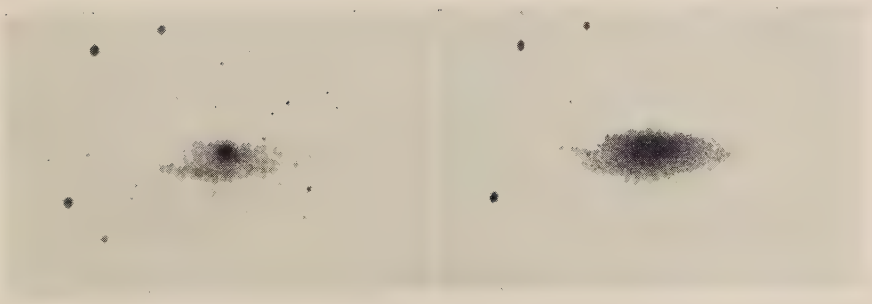
* *Stockholms Observatoriums Annaler*, **14**, No. 3, 1942; *Ap. J.*, **104**, 211, 1946.



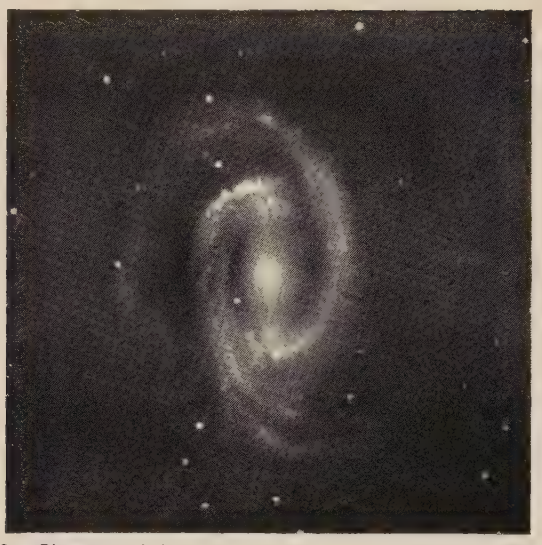
NGC 4565.



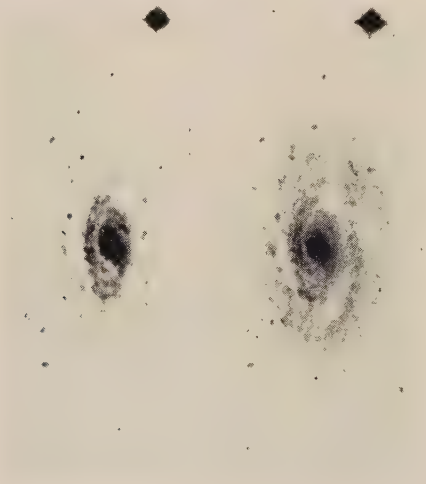
NGC 7331.



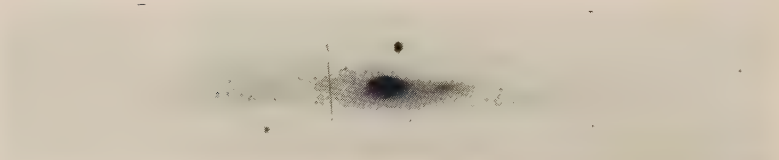
NGC 7331. Central region in ultra-violet (left) and in red (right).
Lick Observatory photographs.



NGC 1300. Photograph by G. Van Biesbroeck, McDonald Observatory.



Central regions of NGC 5055 in blue (left) and in red (right).



NGC 4216. The line across the nebula indicates the path of the magnitude curves in Fig. 10.

of light. In the blue picture the dark clouds of this lane evidently obliterate almost completely a part of a spiral arm very close to the lane on the left-hand side. In the red picture this piece of spiral structure is quite strong compared with its surroundings. This shows that the lane must have the character of a smoke-screen, caused by integration in an enveloping shell of dark clouds about the nucleus, and obstructing a part of the spiral structure on the farther-away side of the screen.

FIG. 10.—*The magnitude curves in blue (full-drawn curve) and in red (dotted curve) along a path crossing the major axis of NGC 4216.*

The nebula NGC 4216 offers another very interesting example of this type of absorption. Fig. 10 gives the results of a photometric tracing in different colours along a path crossing the major axis of the nebula at some distance from the centre, as indicated in the reproduction on Plate 5. The full-drawn curve represents the blue light, the dotted curve the red light. It is very characteristic that on one side of the nebula the colour is nearly uniform, whereas on the other side there is a strong overweight in intensity for the blue light.

This effect may be understood in the following way. We assume schematically that the luminous matter may be divided into two components, two "sub-systems", which have different degrees of flattening towards the equatorial plane. The more flattened component may be assumed to represent the spiral structure. The dark matter, on the other hand, is assumed to possess an intermediate degree of flattening. In the computations of the distribution of light the distributions have been assumed to be spheroidal. The component (1) of the luminous matter has been assumed to have infinite flattening, the ratio of the axes $c/a = 0$, and component (2) to have a ratio of the axes $c/a = 0{\cdot}20$. The total luminosity of component (2) is supposed to be three-tenths of the total luminosity of component (1). The dark matter is supposed to have the ratio of the axes $c/a = 0{\cdot}10$ The optical thickness along the polar axis has been assumed to amount to $0^{m}{\cdot}5$ for

blue light and $0^m.25$ for red light. The lower part of Fig. 11 illustrates the distribution of light for an observer viewing the nebula in such a way that the angle between the line of sight and the nebular plane is about 8°. The full-drawn curve represents the blue light, the dotted curve the red. It is evident that even with the very much simplified schematic assumptions made above the main properties of the observed distribution of colour are quite well reproduced.

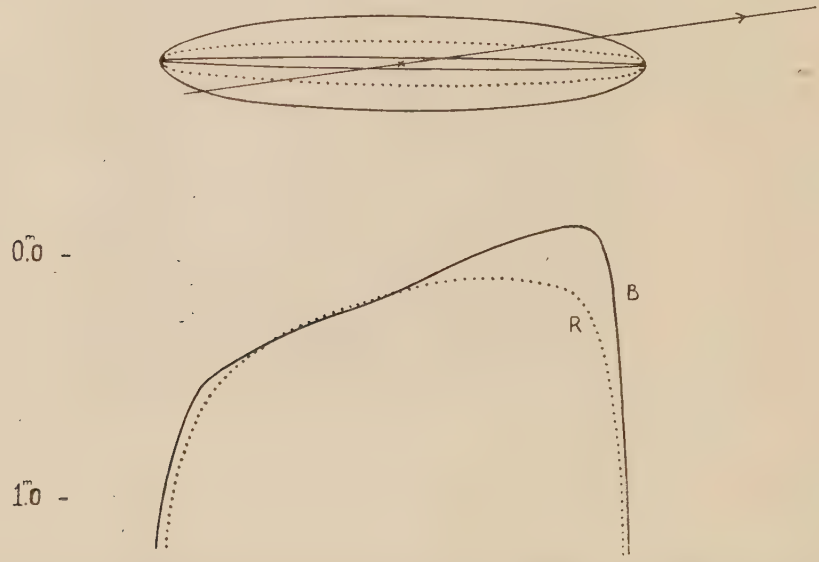


FIG. 11.—*The upper part gives a vertical section of the schematic sub-systems of bright matter (full-drawn ellipses) and of dark matter (dotted ellipse). The lower part gives the corresponding magnitude curves in blue (full-drawn curve) and red (dotted curve).*

In the case of the great Andromeda nebula, M 31, the general absorption is no doubt of the same type as in the nebula just described. The conclusion is that the half of the nebula which is most heavily affected by absorption of light is the more distant part of the nebular disk. On the near part of the disk there is above all a wide absorption region which is apparently centred on the satellite nebula M 32. The interpretation of the absorption effects in M 31 may be checked by a study of the apparent distribution of objects of high intrinsic luminosity. Cepheids, novae and globular clusters. The Cepheids must be expected to be concentrated in the thin layer of the spiral structure; they show a very strong apparent preference for the brighter side of the nebula. The distribution of the novae appears to follow fairly well the general distribution of light intensity in the nebula. A quantitative comparison between the number of novae and the integrated intensity of the background indicates, however, a proportionally greater number of novae in the bright central regions than in the surroundings. This observation, if confirmed, would be of importance for our interpretation of the nova phenomenon, as it would indicate that the density of matter stimulates the process which produces the nova phenomenon. The bright globular clusters (phot. mag. <17) show an essentially different distribution with a wider dispersion over the nebular disk, and on the whole a much less pronounced preference for the brighter side of the nebula. These phenomena may be explained, if the brighter side is the nearer one and if we consider the probable run of the surfaces of constant

optical depth with reference to the observer. If the absorbing matter is not entirely confined to the equatorial plane, these surfaces should have a somewhat smaller inclination towards the line of sight than the equatorial plane itself, and the general obscuring effect on the more distant parts of the nebula for the objects strongly bound to this plane, Cepheids and novae, will be directly accounted for. The apparent distribution of the bright globular clusters may be understood, if these objects may be formally divided up into sub-systems of different degree of flattening towards the equatorial plane. The strongly flattened sub-system will be affected by the absorption much like the stars in general, whereas the less flattened sub-system will be, by projection, more frequent on the more distant, faint side of the nebula. In this way an approximate balance between the two sides as to the total frequency of globular clusters may be established, as actually observed.

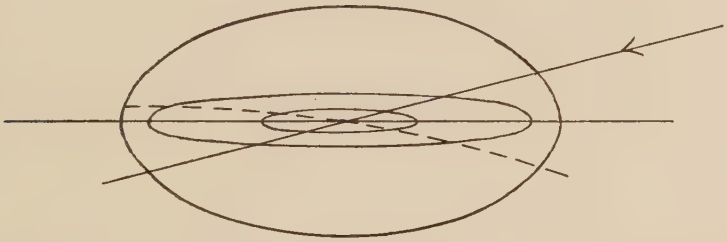


FIG. 12.—*Schematic picture of the distribution of globular clusters in M 31. The dotted curve indicates a surface of constant optical depth along the line of sight.*

We return at last to the dynamical theory. The fairly rapid development of the spiral structure, once it has begun to form, which we have found by comparison between theory and observation for certain objects may seem surprising. It may mean, however, that the development of spiral structure is not strictly a one-way process for one and the same system. It seems probable that the process is quasi-periodic, repeating itself perhaps with certain intermediate periods of rest, during a certain era of the development of the system.

The ultimate cause of the development of spiral structure is no doubt the formation of stars and dust particles out of a primitive gaseous system. This increases the effective degree of flattening towards the plane of rotation, and makes the system unstable against the Riemann wave and the density wave of the "bar". These waves break up the system, as we have seen, either in a "normal" or in a "barred" spiral structure.

In this way the theory of the dynamics of the spiral structure gets an important position in the general theory of cosmical evolution. The theory of the constitution and dynamics of stellar systems, when it once gets a sufficient clarity, as without doubt it can, will form a kind of frame into which the theories of the formation of the elements, and of the generation of the stars, will have to fit. It seems probable that these latter theories, which are at the present time still rather obscure, though promising advances have no doubt been made, may receive here certain guiding principles which on important points may enable us to choose between different possibilities. In this way the theory of the dynamics of the galaxies will get its proper place in the great picture of the general cosmical evolution which it is the ultimate goal of Astronomy to create.

SOLAR MAGNETISM AND THE SUGGESTED FUNDAMENTAL MAGNETIZATION BY ROTATION

S. Chapman

(Received 1948 May 22)

Summary

The magnetic field of the Sun is calculated according to the "fundamental" hypothesis, revived and revised by Blackett, that a rotatory mass flux f produces a magnetic field as if it were a negative electric current of intensity $\beta G^{\frac{1}{2}} f/c$, where β is a factor of order unity, G the gravitational constant and c the speed of light.

The calculation is based on the density distribution given by Blanch, Lowan, Marshak and Bethe for the point convective model of the Sun, and takes into account a non-uniform distribution of the angular velocity ω approximating to that given by Schwarzschild. The corresponding moment of inertia and angular momentum of the Sun are calculated.

On this basis, and taking the Sun's polar magnetic intensity to be 50 (or 10) gauss, β is estimated as 2 (or 0·4) as compared with 0·3 for the Earth.

The inferred intensity at the centre is a few thousand gauss, similar to that of sunspot magnetic fields.

To a depth of about 100,000 km. the field increases downwards approximately as if it were due to a point dipole at the centre. Lower down the increase is less rapid; for the radial component it continues to the centre. The horizontal component attains a maximum about halfway down, where its intensity in the equatorial plane is a few hundred gauss; thence it decreases downwards to zero at about $\frac{1}{3}a$ from the centre (a=solar radius), below which, with reversed direction, it increases steadily to the centre.

The non-uniformity of rotation introduces non-dipole terms into the field potential, which are calculated on the assumption that ω varies linearly with the distance from the axis. Even if ω is doubled or halved at the surface at the equator, compared with its axial value, the non-dipole terms are too small to be observable at present.

Finally it is shown that the magnetic fields of sunspots cannot reasonably be supposed to have this fundamental character. The necessity to attribute these intense fields to some different mechanism (as yet unknown) seems to weaken the case for the "fundamental" hypothesis.

---

1. *Introduction.*—Professor Blackett* has recently reviewed very clearly the facts and theories concerning the general magnetic fields of the Earth, the Sun, and of certain stars.† In particular he emphasized two striking facts, which led him to suggest that these fields indicate a new fundamental property of rotating matter. These facts are that the ratio P/W is (a) nearly the same for these different bodies, and (b) can be very simply expressed as below in terms of fundamental constants: where

P = the axial component of the dipole magnetic moment,

W = the angular momentum;

* P. M. S. Blackett, *Nature*, **159**, 658, 1947. Other theories, not mentioned by Blackett, have been given recently by J. M. Barnóthy, *Hungarian Inst. for Met. and Terr. Mag.*, Paper No. 2, Budapest, 1947; A. Gião, *Gazeta de Matemática*, Nos. 34, 35, 1947/8, Lisbon (and earlier papers there cited); J. Prada Labandera, *Boletin de Meteorologia*, **1**, 1943, and **5**, 11, 1947, Montevideo.

† H. W. Babcock, *Ap. J.*, **105**, 105, 1947; *P.A.S.P.*, **59**, 112, 260, 1947

and

$$P/W = \beta G^{\frac{1}{2}}/2c, \tag{1}$$

where

G = the constant of gravitation,
c = the speed of light,
β = a factor of order unity (for the Earth $\beta = 0{\cdot}30$).

The angular momentum W is equal to

$$\int \rho p^2 \omega \, dv, \tag{2}$$

where ρ denotes the density and ω the angular velocity of the element dv, at the distance p from the axis. If ω is uniform (that is, if the body rotates as if rigid),

$$W = I\omega,$$

where I signifies the moment of inertia about the axis of rotation.

1.1. *The tentative fundamental hypothesis.*—H. A. Wilson* pointed out that for a uniform sphere in rigid rotation, (1) could be explained by supposing that a mass flux ρv, of matter of density ρ moving with velocity v, produces a magnetic field as if it were a negative electric flux whose current intensity C is given by

$$C = K\rho v, \tag{3}$$

where $K = \beta G^{\frac{1}{2}}/c = 8{\cdot}62 \times 10^{-15}\beta$

in c.g.s. electromagnetic units. Wilson proved by experiment that this supposition is not true when v is merely a translatory velocity.

Blackett revived this proposal for the case of *rotatory* motion of a body symmetrical about its axis of rotation. He recognized and discussed the difficulties† of the hypothesis, and of its extension to bodies and motions not symmetrical about an axis.

1.2. *The geomagnetic depth variation.*—At the Royal Society discussion on Blackett's paper, in 1947 May, E. C. Bullard indicated that the fundamental hypothesis would seem to imply a downward *decrease* of intensity of the geomagnetic field within the Earth, instead of the increase generally assumed; and that measurements might check the hypothesis on this point. Measurements since made, in South Africa‡ and in England§, provisionally support the hypothesis, when compared with the calculations of the internal geomagnetic field made by S. K. Runcorn‖ and myself.¶

Another test that may become possible in the future is the measurement of the magnetic field of the Moon.**

1.3. *Further calculations.*—The interest of the "fundamental" hypothesis of Section 1.1 is such that, despite its tentative and incomplete character at present, it is worth while to develop its consequences, particularly as to the nature of the internal fields of the Earth and Sun. This may elicit inferences that will throw light on the validity of the hypothesis; it may also assist in allied theoretical

* H. A. Wilson, *Proc. Roy. Soc.*, A, **104**, 451, 1923.
† See also H. Y. Tzu, *Nature*, **160**, 746, 1947.
‡ A. L. Hales and D. I. Gough, *Nature*, **160**, 746, 1947.
§ S. K. Runcorn, *Nature*, **161**, 462, 1948.
‖ S. K. Runcorn, *Nature*, **160**, 746, 1947; **161**, 52, 1948.
¶ S. Chapman, *Nature*, **161**, 52, 1948; *Ann. de Géophysique*, **4**, 109, 1948.
** S. Chapman, *Nature*, **160**, 395, 1947.

discussions that involve assumptions as to these internal fields—such as theories of sunspot magnetism, or of the secular geomagnetic variation. The fundamental hypothesis, besides having the merit of definiteness, has no strong rival at present.

In my recent paper in the *Annales de Géophysique* I have discussed the internal distribution of the supposed fundamental geomagnetic field, on the basis of a few alternative assumed distributions of density; in this paper I make similar calculations for the Sun. In both cases I treat the rotating body as spherical, with a spherically symmetrical density distribution: that is,

$$\rho = f(r) = \rho(u),$$

where r denotes the distance from the centre O and

$$u = r/a,$$

where a denotes here the solar radius; it is convenient to refer to u as the level.

1.4. *The magnetic moment.*—It is convenient at the outset, however, to notice a simple general theorem deducible from the fundamental hypothesis for any case to which this applies, namely, any case in which there is symmetry of matter and motion about an axis. For non-spherical bodies it is convenient to specify the position of any point P in an axial cross-section by the coordinate z measured parallel to the axis, and the distance p from the axis. Thus in the general symmetrical case, ρ and ω are functions of p and z. A rectangular element of the section, $dp\,dz$, at P, corresponds to a ring round the axis (revolving with speed ωp), and also, according to the fundamental hypothesis, to a circular electric current of strength $K\rho\omega p\,dp\,dz$. This has a magnetic field equal to that of the magnetic shell of this strength, bounded by the circle. The magnetic moment of this shell (equal to its area multiplied by its strength) is

$$\pi K\rho\omega p^3\,dp\,dz,$$

and the magnetic field of the ring current, at a great distance (though not at near points) is that of a dipole of this moment. As the mass and speed of motion of the ring are $2\pi\rho p\,dp\,dz$ and ωp, its angular momentum is $2\pi\rho\omega p^3\,dp\,dz$, so that for each such ring the ratio (dipole moment)/(angular momentum) is $\frac{1}{2}K$. As the dipole moments and angular momenta for all such rings are additive (being unidirectional) it follows that

$$P/W = \tfrac{1}{2}K$$

for the whole body, in accordance with (1). Hence the leading term in the magnetic potential V of the fundamental field at distant outside points is $\frac{1}{2}KW\cos\theta/r^2$, where r, θ are polar coordinates relative to the axis of rotation and any origin O on this axis. In general, there will also be other spherical harmonic terms in V, of higher order; these will depend on the choice of O, the form of the body (that is, of its axial section) and on the nature of the distributions of ρ and ω.

When the body is a uniform sphere in rigid rotation, and O is taken at its centre, the above dipole term constitutes the whole potential V, as H. A. Wilson remarked. This is true also when ρ and ω are non-uniform but are functions of r (or u) only.

2. *The density distribution within the Sun: model A.*—Different workers give decidedly different estimates of the internal density distribution within the Sun: though the total mass (and mean density ρ_m) are known, the distribution can only be inferred from theory. The theoretical results are not expressible

by simple formulae and are given by tables of ρ as a function of u. The divergence of the estimates may be illustrated by the values calculated for the *central density* ρ_c given by Eddington *, Cowling †, and Blanch *et al.* ‡, namely 76, 51 and 110 (taking $\rho = 1$ for water). The value of the density at the centre and higher levels must still be considered rather uncertain, and therefore it would be pointless to be too elaborate and accurate in calculating the internal magnetic field; hence in the calculations of this paper I have mostly been content with slide-rule accuracy. Also it has seemed to me preferable not to make numerical quadratures from the tables of values of ρ, but rather to represent the latter approximately by simple formulae containing suitably chosen parameters, and to use these formulae as the basis of calculation; one can then infer also how a change in the assumed density distribution would affect the inferred magnetic field, by considering other values for the parameters.

The density distribution chosen for this purpose is that given by Blanch *et al.*; a selection from their results, which range from $u=0$ to $u=0{\cdot}833$, is given in Table I, including both $\rho(u)$, the density at the level u, and M_u/M, the fraction of the total mass M that lies *above* the level u. The formulae chosen to represent these values of ρ approximately are

$$\left.\begin{aligned} &\rho/\rho_c = e^{-\alpha u^2} \text{ for } 0 \leqslant u \leqslant \tfrac{1}{2}, \\ &\rho/\rho_c = Ae^{-bu} = A' \times 10^{-b'(u-\frac{1}{2})} \quad \text{for } \tfrac{1}{2} \leqslant u \leqslant 1 \end{aligned}\right\}. \tag{A}$$

The density distribution represented by these formulae, with the values

$$\rho_c = 105{\cdot}5,\ \alpha = 22;\quad \rho_c A' = 0{\cdot}917,\ b' = 6,$$

will be called model A, and is illustrated in Table I for comparison with the values given by Blanch *et al.* The model A distribution has the correct total mass. The two formulae A give different values at their common point ($u=\frac{1}{2}$), but this artificial discontinuity has little influence on the calculated magnetic field, which depends on integrals comparable with that for M_u; the values of M_u/M in columns 3 and 5 of Table I show reasonably good agreement, naturally without any discontinuity at $u=\frac{1}{2}$.

The calculations on model A involve the following integrals:

$$f(b, n, u_1, u_2) \equiv \int_{u_1}^{u_2} u^n e^{-bu}\, du$$

$$= \left[-e^{-bu}\left\{ \frac{u^n}{b} + \frac{nu^{n-1}}{b^2} + \frac{n(n-1)u^{n-2}}{b^3} + \ldots + \frac{n!\,u}{b^n} + \frac{n!}{b^{n+1}} \right\} \right]_{u_1}^{u_2};$$

$$g(\alpha, n, u_1, u_2) \equiv \int_{u_1}^{u_2} u^n e^{-\alpha u^2}\, du;$$

and

$$g(\alpha, 2n+1, u_1, u_2) = \tfrac{1}{2} f(\alpha, n, u_1^2, u_2^2),$$

$$g(\alpha, 2n, u_1, u_2) = \frac{(2n-1)(2n-3)\ldots 5{\cdot}3}{2(2\alpha)^n}\left(\frac{\pi}{\alpha}\right)^{\frac{1}{2}}\left[1 - P(u\sqrt{2\alpha}) \right]_{u_1}^{u_2}$$

$$+ \left[-e^{-\alpha u^2}\left\{ \frac{u^{2n-1}}{2\alpha} + \frac{(2n-1)u^{2n-3}}{(2\alpha)^2} + \ldots + \frac{(2n-1)(2n-3)\ldots 5{\cdot}3u}{(2\alpha)^n} \right\} \right]_{u_1}^{u_2},$$

* A. S. Eddington, *Internal Constitution of the Stars*, Cambridge, 1926, p. 149.
† T. G. Cowling, *M.N.*, **96**, 57, 1935.
‡ G. Blanch, A. N. Lowan, R. E. Marshak and H. A. Bethe, *Ap. J.*, **94**, 37, 1941.

where

$$1-P(x)=\left(\frac{2}{\pi}\right)^{\frac{1}{2}}\int_0^x e^{-\frac{1}{2}t^2}\,dt,$$

which is a tabulated function.

Thus for model A,

$$M=4\pi\int_0^a \rho r^2\,dr=4\pi a^3\int_0^1 \rho u^2\,du=4\pi a^3\rho_c\{g(\alpha,\,2,\,0,\,\tfrac{1}{2})+Af(b,\,2,\,\tfrac{1}{2},\,1)\},$$

whence ρ_m may be obtained, because $M=\frac{4}{3}\pi a^3\rho_m$; also

$$M_u/M=1-(3\rho_c/\rho_m)g(\alpha,\,2,\,0,\,u) \text{ for } 0<u<\tfrac{1}{2},$$
$$M_u/M=(3\rho_c/\rho_m)Af(b,\,2,\,u,\,1) \text{ for } \tfrac{1}{2}<u<1.$$

TABLE I

	Blanch *et al.*		Model A		Model B	
$u=r/a$	ρ	M_u/M	ρ	M_u/M	ρ	M_u/M
0	110	1	105·5	1	28·2	1
0·125	78	0·872	75	0·882	26·4	0·962
0·2	43	0·623	44	0·637	23·7	0·855
0·25	25	0·448	27	0·458	21·2	0·734
0·303	12·6	0·295	14·0	0·285	17·5	0·565
0·370	5·1	0·164	5·1	0·143	12·8	0·321
0·435	2·1	0·082	1·65	0·075	6·8	0·102
0·5	0·917	0·047	0·431 0·917	0·047	0	0
0·625	0·175	0·011	0·163	0·012		
0·714	0·046	0·003	0·048	0·004		
0·833	0·0045	0·000	0·0083	0·001		
0·989*	$1·8\times10^{-6}$*	0·000	$1·1\times10^{-3}$	0·000		
1*	$2·1\times10^{-6}$*	0·000	$9·2\times10^{-4}$	0·000		

* These two values are taken from T. G. Cowling, *M.N.*, **96**, 17, 1935.

2.1. *An alternative density distribution: model B.*—To illustrate how the inferred magnetic field is affected by different degrees of central condensation of the mass, the alternative distribution (model B) given by

$$\rho/\rho_c=1-(u/u_0)^2 \text{ for } 0\leqslant u\leqslant u_0;\quad \rho=0 \text{ for } u>u_0$$

has been considered. This model places all the mass below the level u_0, in effect reducing the radius of the Sun, except that the calculated field will be taken to have the observed surface value at $r=a$. The integrals to be evaluated in this case are of the form

$$F(u_0,\,n,\,\lambda_1,\,\lambda_2)\equiv\int_{\lambda_1 u_0}^{\lambda_2 u_0} u^n\left\{1-\frac{u^2}{u_0^2}\right\}du$$
$$=u_0^{n+1}\int_{\lambda_1}^{\lambda_2}\lambda^n(1-\lambda^2)d\lambda=u_0^{n+1}[\lambda^{n+1}/(n+1)-\lambda^{n+3}/(n+3)]_{\lambda_1}^{\lambda_2}.$$

In this case

$$M_u/M=(3\rho_c/\rho_m)F(u_0,\,2,\,u/u_0,\,1) \text{ for } 0<u<u_0,\ M_u/M=0 \text{ for } u>u_0.$$

To give the mass corresponding to the known value (1·411) of ρ_m,

$$\rho_c=\rho_m/3F(u_0,\,2,\,0,\,1)=5\,\rho_m/2u_0^3.$$

Table I contains values of ρ and M_u/M for model B, taking

$$u_0=\tfrac{1}{2}.$$

The central density is only about $\frac{1}{4}$ that for model A, but the decrease of density, slow near the centre, becomes rapid above $u=\frac{1}{3}$.

2.2. *The moment of inertia of the Sun.*—Denoting this by I,

$$I=2\pi\int_0^a\int_0^\pi \rho r^4 \sin^3\theta\, d\theta\, dr=\frac{8}{3}\pi a^5\int_0^1 \rho u^4\, du.$$

Hence if I_0 is the value of I corresponding to *uniform* density, $\rho=\rho_m$,

$$I_0=\frac{8}{15}\pi\rho_m a^5.$$

Writing $k=I/I_0$, it follows that

$$k=(5/\rho_m)\int_0^1 \rho u^4\, du.$$

For model A,

$$k=(5\rho_c/\rho_m)\{g(\alpha, 4, 0, \tfrac{1}{2})+Af(b, 4, \tfrac{1}{2}, 1)\}$$

and for model B,

$$k=5u_0^2/7.$$

For model A, giving separately the contributions from below and above the level $u=\frac{1}{2}$,

$$k=0\cdot104+0\cdot028=0\cdot132, \qquad (A)$$

and for model B,

$$k=5/28=0\cdot178. \qquad (B)$$

Thus model B has the greater moment of inertia (because of its greater density near $u=\frac{1}{2}$), although ρ_c is only 28 as against 105·5 for model A. In model A the 4·7 per cent of the mass above $u=\frac{1}{2}$ contributes over 20 per cent of the moment of inertia. The value of k for the Sun is unknown, but 0·13 is perhaps likely to be a better estimate of it than the model B value 0·18.

3. *The angular velocity ω within the Sun.*—The equatorial surface angular velocity ω_s is $2\cdot9\times10^{-6}$ radian per second. The surface value at latitude ϕ $(=90^\circ-\theta)$ is given by

$$\omega=\omega_0-\omega_1\cos^2\theta,$$

where $\omega_0=14\cdot4^\circ$ per day, $\omega_1=2\cdot6^\circ$ per day.

The angular velocity within the Sun is uncertain. Some authors (e.g. Jeans) conclude that ω increases with depth; others, such as Schwarzschild*, infer that it decreases downwards. His idealized theory leads to an equatorial surface acceleration as observed, but of twice the true value. His results are illustrated by a diagram showing the lines of equal ω in a meridian section. Except near the axis these lines do not differ greatly from straight lines parallel to the axis, though they bend towards the axis at both ends, thus reducing the surface variation of ω. The convective core, up to the level $u=0\cdot122$, is supposed to rotate with uniform angular velocity ω_c; outside this the line $\omega=\omega_c$ does not meet the Sun's surface, but ends on the axis at $u=0\cdot7$; its greatest distance from the axis is about $0\cdot125a$. The first line of equal ω to reach the axis is approximately $\omega=1\cdot5\omega_c$. The mean distance of the line $\omega=2\omega_c$ from the axis is about $0\cdot4a$, and that of the line $\omega=4\omega_c$ is about $0\cdot85a$.

* M. Schwarzschild, *Ap. J.*, **106**, 427, 1947.

In order to illustrate how the internal fundamental magnetic field would be influenced by a non-uniform distribution of ω, calculations are made for the distribution

$$\omega = \omega_c + \omega' p/a = \omega_c + \omega' u \sin\theta, \qquad \text{(C)}$$

which is a simplified version of Schwarszchild's distribution if ω' is positive, but which with ω' negative will also illustrate a distribution in which ω increases inwards; the lines of equal ω are parallel to the axis. According to this model,

$$\omega_s = \omega_c + \omega'.$$

If $\omega' = 3\omega_c$, (C) would correspond approximately to Schwarzschild's distribution except near the surface; with this model of ω the surface variation of ω is a fourfold increase from pole to equator, but this could be corrected by modifying (C) so that above $u = 0{\cdot}8$ the lines of equal ω bent over towards the axis. This would only very slightly alter the magnetic field above this level.

4. *The angular momentum (W) of the Sun.**—For a sphere in which both ρ and ω may be functions of r and θ,

$$W = 2\pi \int_0^a \int_0^\pi \rho(u, \theta)\omega(u, \theta) r^4 \sin^3\theta \, d\theta \, dr$$

$$= 2\pi a^5 \int_0^1 \int_0^\pi \rho(u, \theta)\omega(u, \theta) u^4 \sin^3\theta \, d\theta \, du.$$

If ρ and ω are functions of u only, this becomes

$$W = \frac{8}{3}\pi a^5 \int_0^1 \rho(u)\omega(u) u^4 \, du$$

or, when ω is uniform, $W = I\omega$.

If ω is given by (C),

$$W = I\omega_c + I'\omega' \equiv I_0(k\omega_c + k'\omega') \equiv k_s I_0 \omega_s,$$

where

$$I' \equiv 2\pi \int_0^a \int_0^\pi \rho r^4 u \sin^4\theta \, d\theta \, dr = \tfrac{3}{4}\pi^2 a^5 \int_0^1 \rho u^5 \, du \equiv k' I_0.$$

For model A it is found that (showing separately the contributions from below and above the level $u = \frac{1}{2}$)

$$\int_0^1 \rho u^5 \, du = 0{\cdot}00905 + 0{\cdot}00484 = 0{\cdot}01389$$

and

$$k' = 0{\cdot}0435.$$

The following table gives the values of $W/I_0\omega_s$ for model A in rotation according to C, for a few typical values of ω'/ω_c.

ω'/ω_c	$-\frac{3}{4}$*	$-\frac{1}{2}$*	0†	1‡	3‡
ω_s/ω_c	$\frac{1}{4}$	$\frac{1}{2}$	1	2	4
$k_s \equiv W/I_0\omega_s$	0·397	0·220	0·132	0·088	0·066

* ω increasing inwards. † Uniform rotation. ‡ ω decreasing inwards.

For model B,

$$\int_0^1 \rho u^5 \, du = \frac{5}{48}\rho_m u_0^3,$$

and for $u_0 = \frac{1}{2}$, $k' = 0{\cdot}0576$, not very different from the value for model A.

* See also page 251.

5. *The dipole moment and β for model A in non-uniform rotation C.*—For the Sun, taking the polar field to be 50 gauss, due to the dipole field only, the dipole moment P is $8{\cdot}5 \times 10^{33}$. Also $I_0\omega_s = 1{\cdot}12 \times 10^{49}$. As $P/W = \frac{1}{2}K$ (Section 1.4), and $W = k_s I_0 \omega_s$, $K = \beta G^{\frac{1}{2}}/c = 8{\cdot}62 \times 10^{-15}\beta$, it follows that $\beta k_s = 0{\cdot}176$. Hence for the above five typical cases of the (C) distribution of ω, β has the following values:

$$\beta = 0{\cdot}44 \qquad 0{\cdot}80 \qquad 1{\cdot}33 \qquad 2{\cdot}0 \qquad 2{\cdot}7$$

In view of the equatorial acceleration of the Sun's surface rotation, and of the arguments given by Schwarzschild for an inward decrease of ω, the most probable estimate of β for the Sun may perhaps be the fourth of these values, 2·0, corresponding to $\omega_s = 2\omega_c$; this is a smaller variation of ω than that deduced by Schwarzschild, but it is to be remembered that he does not take friction into account.

Blackett's estimate of β for the Sun is 1·14; this is based on uniform rotation ($\omega' = 0$, $\omega_s = \omega_c$) and the value $k = 0{\cdot}16$; he also takes a somewhat larger value of P, corresponding to a polar intensity of 53 gauss.*

6. *The electric current distribution and current function J.*—According to the fundamental hypothesis of Section 1.1, the rotatory mass flux acts magnetically as if it were an electric current intensity $C(u, \theta)$ given at the point R, θ, or aU, θ, by

$$C(U, \theta) = K\rho\omega a U \sin\theta.$$

Any thin concentric shell of the sphere, of radius R and small thickness adU, can be regarded as a current sheet with *surface* current intensity $C(U, \theta)adU$. A current function $J(U, \theta)dU$ can be associated with the sheet, such that

$$\frac{\partial J(U, \theta)}{R\partial\theta} = -aC(U, \theta),$$

or

$$\frac{\partial J(U, \theta)}{\partial\theta} = -a^2 UC(U, \theta);$$

and J may be expressed as a series of Legendre functions $P_n(\mu)$, where $\mu = \cos\theta$, namely

$$J(U, \theta) = \sum_{n=0}^{\infty} a_n(U) P_n(\mu),$$

where

$$a_n(U) = \frac{2n+1}{2} \int_{-1}^{1} J(U, \theta) P_n(\mu)\, d\mu.$$

By Rodrigue's formula

$$P_n(\mu) = D^n B_n(\mu),$$

where

$$D \equiv d/d\mu, \quad B_n(\mu) = (\mu^2 - 1)^n / 2^n \,.\, n!$$

Also

$$D^{n-1} B_n = (\mu^2 - 1) D^{n+1} B_n / n(n+1),$$

so that $D^{n-1}B_n$ is zero when $\mu = \pm 1$ or $\theta = 0$, $\theta = \pi$. Hence, integrating by parts,

$$\begin{aligned} a_n(U) &= \frac{2n+1}{2} \int_{-1}^{1} J(U, \theta) D^n B_n(\mu) d\mu = -\frac{2n+1}{2} \int_{-1}^{1} \frac{\partial J}{\partial \mu} D^{n-1} B_n d\mu \\ &= -\frac{2n+1}{2} \int_{\pi}^{0} \frac{\partial J}{\partial \theta} D^{n-1} B_n d\theta = \frac{2n+1}{2} \int_{\pi}^{0} a^2 UC(U, \theta) D^{n-1} B_n \, d\theta \\ &\equiv \frac{2n+1}{2} K\rho(U) a^3 U^2 \alpha_n(U), \end{aligned}$$

* See also page 251.

if ρ is a function of r only; where

$$\begin{aligned}\alpha_n(U) &= \int_{\pi}^{0} \omega(U,\theta)\sin\theta\, D^{n-1}B_n(\mu)\,d\theta \\ &= -\int_{-1}^{1} \omega(U,\theta) D^{n-1}B_n(\mu)\,d\mu \\ &= \frac{1}{n(n+1)}\int_{-1}^{1} \omega(U,\theta)\sin^2\theta\, D^{n+1}B_n(\mu)\,d\mu \\ &= \frac{1}{n(n+1)}\int_{-1}^{1} \omega\sin^2\theta\frac{dP_n(\mu)}{d\mu}\,d\mu \\ &= \frac{1}{n(n+1)}\int_{\pi}^{0} \omega\sin^2\theta\frac{dP_n}{d\theta}\,d\theta \\ &= \frac{1}{n(n+1)}\int_{0}^{\pi} P_n(\mu)\frac{\partial}{\partial\theta}\{\omega(U,\theta)\sin^2\theta\}d\theta,\end{aligned}$$

after another integration by parts.

If ω is an even function of the latitude, the second factor in the last integrand is anti-symmetrical with respect to the equator, and as P_n is symmetrical or anti-symmetrical according as n is even or odd, it follows that

$$\alpha_{2n}=0,\quad a_{2n}=0.$$

If ω is independent of θ,

$$C=K\rho(U)\omega(U)aU\sin\theta,\qquad J=K\rho(U)\omega(U)a^3U^2\cos\theta=a_1(U)P_1(\mu),$$

where $a_1(U)=K\rho(U)\omega(U)a^3U^2$, corresponding to

$$\alpha_1(U)=\tfrac{2}{3}\omega(U),$$

a special case of the above general formula for $\alpha_n(U)$. In this case, which includes that of uniform rotation, J contains only the one Legendre function $P_1(\mu)$, and $a_n=0$ for $n>1$.

6.1. *Non-uniform rotation according to* (C).—If $\omega=\omega_c+\omega' U\sin\theta$, each term contributes independently to a_n and α_n, and it is convenient to write

$$\alpha_n(U)=\omega_c\alpha_{nc}+\omega' U\alpha'_n,$$

where $\alpha_{1c}=\frac{2}{3}$, $\alpha_{nc}=0$ for $n>1$, and

$$\begin{aligned}n(n+1)\alpha'_n &= \int_0^{\pi} P_n(\mu)\frac{d\sin^3\theta}{d\theta}\,d\theta \\ &= \tfrac{3}{4}\int_0^{\pi}(\cos\theta-\cos 3\theta)P_n(\cos\theta)d\theta.\end{aligned}$$

By using the known formulae for P_n as a series of cosines of multiples of θ, this integral can readily be expressed in general terms. The first few values give

$$\alpha'_1=3\pi/16,\quad \alpha'_3=-\pi/128,\quad \alpha'_5=-\pi/2048.$$

7. *The magnetic field.*—The potential dV of the magnetic field of a thin spherical electric current sheet of radius $R(=aU)$, whose current function $J(U,\theta)\,dU$ is expressed as in Section 6 as a series of Legendre functions, is given at a point r, θ (or au, θ) by dV_o at outside points $(u>U)$ and dV_i at inside points

$(u<U)$, where

$$dV_o = 4\pi dU \sum_{n=1}^{\infty} \frac{n}{2n+1} a_n(U) \left(\frac{U}{u}\right)^{n+1} P_n(\mu),$$

$$dV_i = -4\pi dU \Sigma \frac{n+1}{2n+1} a_n(U) \left(\frac{u}{U}\right)^{n} P_n(\mu).$$

Outside the sphere $r=a$, the total field of the whole current distribution of Section 6 has the potential V_o obtained by integrating dV_o from $U=0$ to $V=1$, namely

$$V_o = 2\pi K a^3 \Sigma n P_n(\mu) f_n(1)/u^{n+1},$$

where

$$f_n(u) = \int_0^u U^{n+3} \rho(U) \alpha_n(U) dU,$$

so that

$$f_n(0) = 0.$$

The first term in V_o is the potential of a dipole at O, directed along the axis, with dipole moment P given by

$$P = 2\pi K a^5 f_1(1) = 2\pi K a^5 \int_0^1 U^4 \rho(U) \alpha_1(U) dU$$

$$= \pi K a^5 \int_0^1 \int_0^\pi U^4 \rho(U) \omega(U, \theta) \sin^3\theta \, d\theta \, dU = \tfrac{1}{2} K W;$$

this is merely a special case (for the sphere) of the general theorem of Section 1.4. When ω is independent of θ, so that $\alpha_n = 0$ for $n>1$, V_o has only this one term, and the field is a purely dipole field.

Inside the sphere $r=a$, the field has no scalar potential, though it can be expressed as the curl of a vector potential. Its radial and horizontal components B_r, B_θ (in the directions of r and θ increasing) can be obtained without evaluating the vector potential, by integrating the contributions made by each spherical shell.

The shell considered above has a field whose components dB_r, dB_θ $(= -\partial dV/\partial r, \; -\partial dV/r\partial\theta)$ are given at a point au, θ outside the shell $(u>U)$ by

$$dB_{ro} = -4\pi dU \Sigma \frac{n(n+1)}{2n+1} \frac{a_n(U)}{au} \left(\frac{U}{u}\right)^{n+1} P_n(\mu),$$

$$dB_{\theta o} = 4\pi dU \Sigma \frac{n}{2n+1} \frac{a_n(U)}{au} \left(\frac{U}{u}\right)^{n+1} \frac{dP_n}{d\theta},$$

and inside the shell $(u<U)$ by

$$dB_{ri} = -4\pi dU \Sigma \frac{n(n+1)}{2n+1} \frac{a_n(U)}{au} \left(\frac{u}{U}\right)^{n} P_n(\mu),$$

$$dB_{\theta i} = -4\pi dU \Sigma \frac{n+1}{2n+1} \frac{a_n(U)}{au} \left(\frac{u}{U}\right)^{n} \frac{dP_n}{d\theta}.$$

To get $B_r(u, \theta)$ and $B_\theta(u, \theta)$ at a point au, θ within the sphere $(u<1)$, it is necessary to integrate dB_{ro} and $dB_{\theta o}$ from $U=0$ to $U=u$, and add the integral of dB_{ri} or $dB_{\theta i}$ from $U=u$ to $U=1$. Thus, after substituting for $a_n(U)$ in terms of $\alpha_n(U)$ from Section 6,

$$B_r(u, \theta) = -2\pi K a^2 \Sigma n(n+1) P_n(\mu) \{ f_n(u)/u^{n+2} + u^{n-1} g_n(u) \},$$

$$B_\theta(u, \theta) = 2\pi K a^2 \Sigma (dP_n/d\theta) \{ n f_n(u)/u^{n+2} - (n+1) u^{n-1} g_n(u) \},$$

where $f_n(u)$ is as already defined, and

$$g_n(u) = \int_u^1 U^{2-n}\rho(U)\alpha_n(U)dU, \quad g_n(1) = 0.$$

8. *The internal field: ω independent of θ.*—In this case $a_n = 0$ for $n > 1$, and therefore $f_n = 0$, $g_n = 0$ for $n > 1$ (Section 6). The internal field, like the field outside, involves only the one Legendre function $P_1(\cos\theta)$, and the dependence of B_r and B_θ on u and on θ is represented by separate factors, as in

$$B_r(u, \theta) = -\frac{8}{3}\pi K a^2 \cos\theta\{f_1(u)/u^3 + g_1(u)\},$$

$$B_\theta(u, \theta) = -\frac{4}{3}\pi K a^2 \sin\theta\{f_1(u)/u^3 - 2g_1(u)\}.$$

Hence the *proportionate* variation of B_r and B_θ is the same along all radii (θ constant) or over any concentric sphere (u constant). In the latter case each component varies with θ as for a dipole field (because B_r varies as $\cos\theta$ and B_θ as $\sin\theta$) but this is not so for the vector B as a whole, because here the ratio $(B_r/\cos\theta) : (B_\theta/\sin\theta)$ is not 2 as for a dipole field.

It is convenient to indicate the depth-distribution of the field by the ratios $b_r(u)$, $b_\theta(u)$ defined by

$$b_r(u) \equiv B_r(u, \theta)/B_r(1, \theta), \quad b_\theta(u) = B_\theta(u, \theta)/B_\theta(1, \theta),$$

that is, the ratios of B_r and B_θ to their surface values on the same radius. Thus

$$b_r(u) = \{f_1(u)/u^3 + g_1(u)\}/f_1(1),$$

$$b_\theta(u) = \{f_1(u)/u^3 - 2g_1(u)\}/f_1(1).$$

The central magnetic induction B_c is along the axis and is given by

$$B_c = B_r(0, 0) = -B_\theta(0, \tfrac{1}{2}\pi) = -8\pi K a^2 g_1(0)/3.$$

8.1. *The internal magnetic field: uniform rotation: models A, A′, B.*—The integrals f_n and g_n have been calculated for models A and B in *uniform* rotation, using the formulae of Section 2, with the values of the parameters there indicated.

Similar calculations have been made also for an allied model density distribution (A′) given by

$$\rho = \rho_c\, e^{-\alpha u^2} \quad 0 \leqslant u \leqslant 1,$$

taking $\alpha = 22$ and ρ_c as 109·7 (almost the same as the value given by Blanch *et al.*); these parameters give the correct total mass. The values of ρ for model A′ are the same, over the range $0 \leqslant u \leqslant \frac{1}{2}$, as those for model A in Table I, increased in the ratio 109·7/105·5; above $u = \frac{1}{2}$ the model A′ has only 0·77 per cent of its total mass, as against 4·7 per cent for model A and the model of Blanch *et al.*; at $u = 0{\cdot}625$ model A′ gives $\rho = 0{\cdot}0621$ instead of 0·175 as given by Blanch *et al.*, and above this level ρ continues to decrease, in model A′, much too rapidly.

From the values of f_n and g_n thus calculated, b_r and b_θ have been determined, and are given in the table below. In the case of model A the values of b_i are also given, where $b_i \equiv f_1(u)/u^3 f_1(1)$; it is the part of b_r and b_θ due to the part of the sphere *below* the level u. The last line of the table gives $1/u^3$, which would be the value of b_r and b_θ if the field were due to a point dipole at O.

The internal field : uniform rotation

$u=$	0	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{2}$	$\frac{5}{8}$	$\frac{3}{4}$	$\frac{7}{8}$	1
				Model A					
$b_r=$	65	54	31	13	7·3	4·0	2·3	1·5	1
$b_\theta=$	−131	− 86	−20	2·9	4·2	3·3	2·3	1·5	1
$b_i=$	0	7	14	9·4	6·3	3·8	2·3	1·5	1
				Model A′					
$b_r=$	78	63	36	15	7·6	4·1	2·4	1·5	1
$b_\theta=$	−155	−101	−22	4·2	7·5	4·0	2·4	1·5	1
				Model B					
$r=$	35	32	25	16	8	4·1	2·4	1·5	1
$b_\theta=$	− 70	− 60	−34	4	8	4·1	2·4	1·5	1
				Point dipole					
$1/u^3=$	∞	512	64	19	8	4·1	2·4	1·5	1

Among the points worthy of note in this table the following may be mentioned:

(1) Above the level $u=\frac{3}{4}$ (that is, in the outer layer down to a depth of 175,000 km.) the field is almost identical with the point dipole field, and proceeds almost wholly from below this level; this outer layer contains too little mass to make a significant contribution to the field. Even down to the level $u=\frac{1}{2}$ the departure from the dipole field is only moderate, at least as regards B_r.

(2) b_r (or B_r) steadily increases downwards to the centre; below $u=\frac{1}{4}$ the departure from the dipole field increases greatly.

(3) b_θ (or B_θ) at first increases downwards (whereas for the Earth it decreases downwards right from the surface, because the terrestrial density does not decrease very greatly towards the surface). It rises to a moderate maximum, which in the table ranges from about 4 to 8, corresponding in the equatorial plane (where $B_r=0$) to an intensity of from about 100 to about 200 gauss, at a level round about $u=\frac{1}{2}$.

(4) Below this level of positive maximum, b_θ decreases to zero at about the level $u=\frac{1}{3}$; in the case of the Earth the corresponding level is much higher, about $u=\frac{7}{8}$.

(5) In my paper (*Annales de Géophysique*) on the Earth's field, I show that the lines of force have the equation

$$u^2 b_r(u) \sin^2\theta = C^2,$$

where C denotes a parameter defining a particular line; this equation is valid whenever ρ and ω are functions of r only, independent of θ. The lines are of oval form (except the special line $C=0$, which is the axis of rotation), encircling their limiting point or "focus", which (in any axial section) is in the equatorial plane at the level where b_θ is zero. Thus the focus is far deeper in the Sun than in the Earth, owing to the much greater central condensation of mass; in the Sun all the lines of force pass through the circle in the equatorial plane, of radius about $\frac{1}{3}a$, containing the points at which $B_r=B_\theta=0$.

(6) Below the level of the focus B_θ is reversed in direction and steadily increases (numerically) towards the centre.

(7) The internal contribution (b_i) to b_r and b_θ increases upwards from zero at the centre to a maximum, which for model A is about 14·5 at about the level $u=0{\cdot}2$; above the level of maximum it declines steadily to its surface value 1.

(8) The central intensity in the case of models A, A′ and B is 65, 78 and 35 times the polar surface intensity, that is, 3250, 3900 and 1750 gauss. These intensities are similar, in order of magnitude, to sunspot magnetic intensities.

Of the three models A, A′ and B, the order of decreasing central density is A′ (109·7), A (105·5), B (28), which is also the order of decreasing central field intensity ($b_r = 78, 65, 35$); but the latter is by no means proportional to ρ_c. If the density were uniform ($\rho_c = \rho_m = 1{\cdot}411$) the central value of b_r would be 2·5. The considerable difference between the results for models A and A′ is somewhat surprising, and suggests that rather greater values of $b_r(0)$ might be given by a model having somewhat greater central condensation than model A′.

9. *Non-uniform rotation C; the non-dipole component of the surface field, model A.*—Non-uniform rotation, if ω varies with θ, introduces higher (non-dipole) harmonic terms in the potential. This will here be considered only for the model A, rotation C, at the surface. The preceding formulae lead to the following expressions for B_r and B_θ at the surface, as far as the harmonic

$$P_3 = \tfrac{1}{2}(5\cos^2\theta - 3)\cos\theta,$$

namely

$$B_r(1, \theta) = -4\pi K a^2 \rho_m \{0{\cdot}023\omega_c P_1 - 0{\cdot}00033\omega' P_3\},$$

$$B_\theta(1, \theta) = 4\pi K a^2 \rho_m \{0{\cdot}012\omega_c dP_1/d\theta - 0{\cdot}00008\omega' dP_3/d\theta\}.$$

If ω' has the same sign as ω_c (that is, if the angular velocity increases with distance from the axis, as inferred by Schwarzschild and in conformity with the surface equatorial acceleration), the terms P_1 and P_3 reinforce one another at the equator, where $B_r = 0$, and oppose one another at the poles, where $B_\theta = 0$. At the poles $P_1 = P_3 = \pm 1$, and the ratio of the P_3 to the P_1 contribution to B is $-0{\cdot}014\,\omega'/\omega_c$, namely $-1/70$ if $\omega' = \omega_c$, or $-1/23$ if $\omega' = 3\omega_c$; at the equator $dP_1/d\theta = -\sin\theta = -1$, $dP_3/d\theta = \frac{1}{2}\sin\theta(3 - 5\cos^2\theta) = \frac{3}{2}$, and the ratio of the P_1 to the P_3 contribution to B is $+0{\cdot}0105\,\omega'/\omega_c$. These departures from the dipole field are too small to be observable on the Sun with present techniques, if ω'/ω_c (whatever its sign) does not exceed 3. The contribution of the higher harmonics P_5, P_7 seems to be still less important.

It is of interest to note that though in model A 95 per cent of the mass is below the level $u = \frac{1}{2}$, two-thirds of the surface P_3 field comes from above this level. The rotation distribution C would seem to exaggerate the variation of ω within the layer above $u = \frac{1}{2}$ (unless ω'/ω_c is taken much less than 1), judging from the relatively small surface variation of ω.

10. *Sunspot magnetism.*—Until H. W. Babcock succeeded recently in measuring stellar magnetic fields of intensity up to about 8000 gauss, the most intense large-scale magnetic fields known were those of sunspots, whose intensities range up to about 4000 gauss. It is natural to enquire whether these can be explained as "fundamental" fields due to rotation.

The collective properties of sunspots and their magnetic polarities have not been explained, but qualitatively the best available theory of them is that given by V. Bjerknes.* He postulates two subsurface vortices in each hemisphere (northern and southern), surrounding the axis; in alternate sunspot cycles one is uppermost in each hemisphere, and is moving slowly to lower

* V. Bjerknes, *Ap. J.*, **64**, 93, 1926.

latitudes, while the deeper vortices move polewards to rise to the surface at the next sunspot minimum. The vortices of each pair have opposite senses of internal rotation. The appearance of a sunspot is attributed to the rise of one of the upper vortices to the surface, where its broken ends, showing opposite rotations, represent the two spots of a bipolar pair. Below the surface the cross-section of the vortices is supposed to be far from circular, the vertical thickness being much less than the horizontal breadth (along the meridian).

An alternative theory is that the two spots of a bipolar pair are connected by a subsurface vortex extending from one to the other, forming half an oval. This theory does not explain why these isolated vortices show the known remarkable collective properties of sunspots.

This difference is however immaterial to the purpose of examining whether any such vortex motion could give rise to a "fundamental" magnetic field due to the rotation, of intensity comparable to 3000 gauss. For this purpose it is legitimate to disregard the curvature of the vortex axis, and to treat its section as circular: and even to ignore the variation of angular velocity round the axis, at different distances from it.

Hence I consider an infinitely long straight circular cylindrical vortex of radius a in rotation about its axis with uniform angular velocity ω. At distance r from the axis the mass flux is $\rho\omega r$ and the equivalent electric intensity is $K\rho\omega r$; this is equal to $-(\frac{1}{4}\pi)dB/dr$; where B denotes the magnetic intensity at this distance, parallel to the axis. Hence on the axis the magnetic intensity B_0 is given by

$$B_0 = 4\pi \int_0^a K\rho\omega r\, dr = 2\pi K\rho\omega a^2,$$

if ρ and ω are treated as constant. The total magnetic flux along the vortex is

$$F = 4\pi^2 K \int_0^a \rho\omega r^3\, dr = \pi^2 K\rho\omega a^4 = \tfrac{1}{2}\pi a^2 B_0,$$

if ρ and ω are treated as constant, though in fact both would vary with r.

The value of F for a great sunspot may be taken as 10^{22} gauss cm.^{2}, though this value has been exceeded fourfold.*

The depth of the vortex is unknown, but seems unlikely to exceed the maximum separation between the two spots of great bipolar pairs; it may well be much less. This distance is about 10° of longitude, or about 100,000 km. At this depth ($u = 0{\cdot}86$) the density ρ, according to Blanch *et al.*, will be less than 0·005 or 5×10^{-3}.

The mean radius a of the vortex must be supposed to be only a fraction of the depth, say 20,000 km. This is the radius of the visible area of a great spot (though it has sometimes been exceeded). I have suggested† that the vortex may "funnel out" at the Sun's surface, and this seems to me likely on hydrodynamical grounds, though doubt has been cast by Cowling‡ on the electromagnetic argument I advanced in support of this view.

These values of F, ρ and a may be inserted in the relation $F = \pi^2 K\rho\omega a^4$ to obtain the value of $\beta\omega$ necessary to explain the sunspot field as a fundamental field due solely to the rotation. The result is

$$\beta\omega = 1{\cdot}5 \text{ radians per second.}$$

* S. Chapman, *M.N.*, **103**, 127, 1943.
† *Loc. cit.*
‡ T. G. Cowling, *M.N.*, **106**, 222, 1946.

Even if β is taken as large as 10 (instead of 0·3, the value for the Earth, the only body for which it has been determined with certainty), the absurdly excessive value 0·15 radian per second (or one revolution in 42 seconds) is obtained for ω. The assumed values of ρ and a are altogether too small to give a fundamental sunspot field of the observed intensity, and it does not seem possible to increase either or both sufficiently for the purpose.

As Professor Blackett has remarked to me, Sutherland * in 1908 came to a somewhat similar conclusion, by an argument of which many details would today need revision.

11. *General discussion.*—In 1918 Hale *et al.* † announced that the general magnetic field discovered by Hale ‡ in 1913 had the remarkable property of radial limitation, that is, its intensity was found to decrease to the lower limit of measurement (about 10 gauss) in about 300 km. If true, the field must differ greatly from a dipole field; as Rosseland § and I ‖ pointed out, the lines of force must be nearly horizontal. The limitation would also imply the presence of electric currents in the Sun's atmosphere.

In 1928 I proposed a theory ¶ to explain the existence of such electric currents, but later withdrew it in agreement with criticisms by Cowling **; this implied also the withdrawal of my inference, in the same paper, that the downward increase of the field continued below the lowest level of observation, so that an intensity of 10,000 gauss was estimated at the moderate depth of 350 km.

Nevertheless, if the evidence for the radial limitation is accepted, it is not unreasonable to suppose, merely on the ground of continuity, that the rapid downward increase of the field continues for some distance below the surface. In a discussion of sunspot magnetism in 1933, Cowling †† considered "the possibility that the sunspot field may arise as a disturbance in the (almost horizontal) general field of the Sun. . . . For the present hypothesis to be correct, the general field must possess an intensity at lower levels not less than, say, 1000 gauss, a value which, in the absence of data to the contrary, cannot be regarded as impossible" (*loc. cit.*, pp. 43, 46).

It should be remembered, however, that the lack of further confirmation of the radial limitations of the Sun's general magnetic field has caused doubts to grow as to its reality, though Thiessen ‡‡ has confirmed the existence and magnitude of the field itself.

In the absence of any better guidance, it may be noted that if the Sun's general magnetic field has the fundamental character suggested by Blackett, a theory of sunspot magnetic fields such as that proposed by Cowling in 1933 would require the convection of a bundle of tubes of magnetic force from a very great depth, e.g. 400,000 or 500,000 km., unless the bundle can become more concentrated as it rises.

* W. Sutherland, *Terr. Mag.*, **13**, 155, 1908.
† G. E. Hale, F. H. Seares, A. v Maanen, and F. Ellerman, *Ap. J.*, **47**, 206, 1918.
‡ G. E. Hale, *Ap. J.*, **38**, 27, 1913.
§ S. Rosseland, *Ap. J.*, **62**, 387, 1925.
‖ S. Chapman, *M.N.*, **89**, 57, 80, 1928.
¶ *Loc. cit.*
** T. G. Cowling, *M.N.*, **90**, 140, 1929.
†† T. G. Cowling, *M.N.*, **94**, 39, 1933.
‡‡ G. Thiessen, *Ann. de Géophysique*, **9**, 101, 1946.

Such a theory has indeed been advocated by Alfvén*, but I share the preference expressed by Cowling† in a more recent discussion of sunspot fields, for a theory like that of Bjerknes in which these fields are regarded as local breaches in a girdle of tubes of force, distinct from the general field, and normally encircling the Sun's axis beneath the surface.

If the vortex tube and its magnetic bundle funnels out near the surface, as I suggested in a presidential address‡, the bundle of magnetic tubes below the surface may be more intense than the general magnetic field even at its strongest, at the centre.

The existence of the intense sunspot magnetic fields, which the fundamental theory (in anything like its present form) seems incapable of explaining, seems to weaken the case for postulating a new fundamental property of matter to explain the general magnetic field; the more so as the present discussion, based on what seems the best available evidence as to the density and rotation within the Sun, gives a value for β several times as large as that for the Earth.

Nevertheless it remains remarkable, and may be profoundly significant, as Blackett urges, that the ratios of the values of P, extending over a considerable range, for the Earth, the Sun, and certain stars, to the associated values of W, differ by less than one factor of ten.

Addendum, 1948 *October* 9

Thiessen has recently withdrawn his confirmation of the Mount Wilson estimated value 50 gauss for the Sun's surface magnetic intensity at the poles. He now considers that it cannot exceed 10 gauss. The corresponding value of β for the Sun as estimated in Section 5 must therefore be reduced from 2·0 (taking the value there considered as possibly the most likely one) to not more than 0·4. This upper limit is close to the value (0·30) of β for the Earth.

S. Lundquist§ has lately estimated the angular momentum W of the Sun, using the density model of Blanch *et al.*, as here, but basing his estimate of ω on a theorem of Ferraro‖ in association with two different "models" of the lines of magnetic force within the Sun, by Alfvén¶ and Cowling.** For the former, Lundquist estimates W as $1{\cdot}7 \times 10^{48}$; he states that for Cowling's model it would be "somewhat lower". My estimates of W, based on a (modified) Schwarzschild distribution of ω, are $1{\cdot}0 \times 10^{48}$ and $0{\cdot}7 \times 10^{48}$, according as ω'/ω_c is taken as 1 or 3.

Queen's College, Oxford:
1948 *May* 21.

* H. Alfvén, *Arkiv f. Mat., Astr. och Fys.*, **29**B, No. 2, 1942; **29**A, No. 11, 1943.
† T. G. Cowling, *M.N.*, **106**, 224, 1946.
‡ S. Chapman, *M.N.*, **103**, 127, 1943.
§ S. Lundquist, *Arkiv f. Mat., Astr. och Fys.*, **35**A, No. 27, 1948.
‖ V. C. A. Ferraro, *M.N.*, **97**, 458, 1937.
¶ H. Alfvén, *M.N.*, **105**, 388, 1945.
** T. G. Cowling, *M.N.*, **105**, 166, 1945.

THE STEADY-STATE THEORY OF THE EXPANDING UNIVERSE

H. Bondi and T. Gold

(Received 1948 July 14)

Summary

The applicability of the laws of terrestrial physics to cosmology is examined critically. It is found that terrestrial physics can be used unambiguously only in a stationary homogeneous universe. Therefore a strict logical basis for cosmology exists only in such a universe. The implications of assuming these properties are investigated.

Considerations of local thermodynamics show as clearly as astronomical observations that the universe must be expanding. Hence, there must be continuous creation of matter in space at a rate which is, however, far too low for direct observation. The observable properties of such an expanding stationary homogeneous universe are obtained, and all the observational tests are found to give good agreement.

The physical properties of the creation process are considered in some detail, and the possible formulation of a field theory is critically discussed.

1. *The perfect cosmological principle*

1.1. The unrestricted repeatability of all experiments is the fundamental axiom of physical science. This implies that the outcome of an experiment is not affected by the position and the time at which it is carried out. A system of cosmology must be principally concerned with this fundamental assumption and, in turn, a suitable cosmology is required for its justification. In laboratory physics we have become accustomed to distinguish between conditions which can be varied at will and the inherent laws which are immutable.

Such a distinction between the "accidental" conditions and the "inherent" laws and constants of nature is justifiable so long as we have control over the "accidental", and can test the validity of the distinction by a further experiment. In astronomical observations we do not have this control, and we can hence never prove which is "accidental" and which "inherent". This difficulty, though logically a very real one, need not concern us in an interpretation of the dynamics of the solar system. We may be satisfied when we discover that the solar system with all its numerous orbits is accurately one of the many systems permitted by our "inherent" laws.

But when we wish to consider the behaviour of the entire universe, then the logical basis for a distinction between "inherent" laws and "accidental" conditions disappears. Any observation of the structure of the universe will give as unique a result as, for instance, a determination of the velocity of light or the constant of gravitation. And yet, if we were to contemplate a changing universe we should have to assume some such observations to represent "accidental" conditions and others "inherent" laws.

Such assumptions were in fact implied in all theories of evolution of the universe; they were necessary to specify the problem. Without them, there would be no rules and hence unlimited freedom in any extrapolation into the future or into the past. Some such sets of assumptions may be intellectually much more

satisfying than others, and accordingly they are adopted. We may place much reliance in such aesthetic judgments, but we cannot claim any logical foundation for them.

1.2. Let us take an example to demonstrate the point of this criticism. Present observations indicate that the universe is expanding. This suggests that the mean density in the past has been greater than it is now. If we are now to make any statement regarding the behaviour of such a denser universe, and the way in which it is supposed to have reached its present condition, then we have to know the physical laws and constants applicable in denser universes. But we have no determinations for those. An assumption which is commonly made is to say that those physical laws which we have learnt to regard as "intrinsic" because they are unaffected by any changes of conditions which we may produce, are in fact not capable of any change. It is admitted that such a change of the density of the universe would have slight local effects; there would be more or less light arriving here from distant galaxies. But, it is assumed, there would be no change in the physical laws which an observer would deduce in a laboratory.

Such a philosophy may be intellectually very agreeable : it gives permanence to the abstract things, the physical laws, whilst it regards the present condition of the universe as merely a particular demonstration of the consequences of such laws. There are, however, grave difficulties inherent in such a view. The most striking of those is concerned with the absolute state of rotation of a body. Mach (1) examined this problem very thoroughly and all the advances in theory which have been made have not weakened the force of his argument. According to "Mach's Principle" inertia is an influence exerted by the aggregate of distant matter which determines the state of motion of the local frame of reference by means of which rotation or acceleration is measured. A particular rotational state which is described as "non-rotating" can be found by a purely local experiment (a Foucault pendulum, for instance). This rotational state is always found to coincide accurately with the rotational state of the distant stars around the observer. Mach's principle is the recognition that this coincidence must in fact be due to a causal relation. If, as is now widely agreed, we adopt Mach's principle, then we imply that the nature of any local dynamical experiment is fundamentally affected by distant matter. We can hence not contemplate a laboratory which is shielded so as to exclude all influence from outside; and for the same reason we cannot have any logical basis for choosing physical laws and constants and assigning to them an existence independent of the structure of the universe.

Any interdependence of physical laws and large-scale structure of the universe might lead to a fundamental difficulty in interpreting observations of light emitted by distant objects. For if the universe, as seen from those objects, presented a different appearance, then we should not be justified in assuming familiar processes to be responsible for the emission of the light which we analyse. This difficulty is partly removed by the "cosmological principle". According to this principle all large-scale averages of quantities derived from astronomical observations (i. e. determination of the mean density of space, average size of galaxies, ratio of condensed to uncondensed matter, etc.) would tend statistically to a similar value independent of the positions of the observer, as the range of the observation is increased; provided only that the observations from different places are carried out at equivalent times. This principle would mean that there is nothing outstanding about any place in the universe, and that those differences which do

exist are only of local significance; that seen on a large scale the universe is homogeneous.

This principle is widely recognized, and the observations of distant nebulae have contributed much evidence in its favour. An analysis of these observations indicates that the region surveyed is large enough to show us a fair sample of the universe, and that this sample is homogeneous.

But in the sense in which the principle came to be adopted there is the qualification regarding the time of the observation. The universe is still presumed to be capable of altering its large-scale structure, but only in such a way as not to upset its homogeneity. The result of a large-scale observation may hence be a measure of a universal time.

We might have looked to the cosmological principle for a justification of the assumption of the general validity of physical laws; but whilst the principle supplies the justification with respect to changes of place, it still leaves the possibility of a change of physical laws with universal time. Any system of cosmology must still involve a speculation about this dependence as one of its basic assumptions. Indeed, we are not even in a position to interpret observations of very distant objects without such an assumption, for the light which we receive from them was emitted at a different instant in this scale of universal time, and accordingly the processes responsible for its emission may be unfamiliar to us.

The systems of cosmology may well be classified according to the assumption made or implied at this stage of the argument. While one school of thought considers all the results of laboratory physics to be always applicable, without regard to the state of the universe, another starts from the narrow cosmological principle and with the aid of a number of intellectually agreeable assumptions arrives at the conclusion that laboratory physics is qualitatively permanently valid, though some of its "constants" are changing. There are yet other schools of thought which attempt to distinguish the changeable from the permanent constants by their magnitudes.

We shall proceed quite differently at this point. As the physical laws cannot be assumed to be independent of the structure of the universe, and as conversely the structure of the universe depends upon the physical laws, it follows that there may be a stable position. We shall pursue the possibility that the universe is in such a stable, self-perpetuating state, without making any assumptions regarding the particular features which lead to this stability. We regard the reasons for pursuing this possibility as very compelling, for it is only in such a universe that there is any basis for the assumption that the laws of physics are constant; and without such an assumption our knowledge, derived virtually at one instant of time, must be quite inadequate for an interpretation of the universe and the dependence of its laws on its structure, and hence inadequate for any extrapolation into the future or the past.

Our course is therefore defined not only by the usual cosmological principle but by that extension of it which is obtained on assuming the universe to be not only homogeneous but also unchanging on the large scale. This combination of the usual cosmological principle and the stationary postulate we shall call the *perfect cosmological principle*, and all our arguments will be based on it. The universe is postulated to be homogeneous and stationary in its large-scale appearance as well as in its physical laws.

We do not claim that this principle must be true, but we say that if it does not hold, one's choice of the variability of the physical laws becomes so wide that cosmology is no longer a science. One can then no longer use laboratory physics without relying on some arbitrary principle for their extrapolation.

But if the perfect cosmological principle is satisfied in our universe then we can base ourselves confidently on the permanent validity of all our experiments and observations and explore the consequences of the principle. Unless and until any disagreement appears we therefore accept the principle, since this is the only assumption on the basis of which progress is possible without further hypothesis. As will be seen later many conclusions can be drawn which can be compared with experiment. It would be most fortuitous if no contradictions arose, although the perfect cosmological principle was not valid.

2. *Application of the perfect cosmological principle*

2.1. Although the perfect cosmological principle states that the universe is homogeneous and stationary on the large scale no statement is made about spatial isotropy. This is usually assumed together with homogeneity for mathematical convenience, but in our view the position of the two assumptions is quite different. Homogeneity and the stationary property are vital to the very existence of the subject. There may be similarly strong arguments for isotropy, but those so far advanced do not seem to us to be equally compelling. Further considerations of the nature of the stable self-perpetuating state referred to above can be made, which suggest but do not prove isotropy.

We shall therefore introduce spatial isotropy as an assumption which is strongly suggested both by local and by astronomical observations and which is simplest.

We must now apply the perfect cosmological principle to laboratory physics and to astronomical observations. We regard the principle as of such fundamental importance that we shall be willing if necessary to reject theoretical extrapolations from experimental results if they conflict with the perfect cosmological principle even if the theories concerned are generally accepted. Of course we shall never disregard any direct observational or experimental evidence and we shall see that we can easily satisfy all such requirements.

2.2. For the perfect cosmological principle to apply, one might at first sight expect that the universe would have to be static, i. e. to possess no consistent large-scale motion. This, however, would conflict with the observations of distant galaxies, and it would also conflict with the thermodynamic state which we observe. For such a static universe would be very different indeed from the universe we know. A static universe would clearly reach thermodynamical equilibrium after some time. An infinitely old universe would certainly be in this state. There would be complete equilibrium between matter and radiation, and (apart possibly from some slight variations due to gravitational potentials) everything would be at one and the same temperature. There would be no evolution, no distinguishing features, no recognizable direction of time. That our universe is not of this type is clear not only from astronomical observations but from local physics and indeed from our very existence. Accordingly there must be large-scale motions in our universe. The perfect cosmological principle permits only two types of motion, viz. large-scale expansion with a velocity proportional to distance, and its reverse, large-scale contraction.

In a contracting universe there would be even more radiation compared with matter than in a static universe. Therefore we reject this possibility and confine our attention to an expanding universe.

2.3. The observations of distant galaxies, which are now capable of a more rigorous interpretation by means of the perfect cosmological principle, inform us of the motion of expansion. This motion in which the velocity is proportional to the distance (apart from a statistical scatter) is well known to be of the only type compatible with homogeneity; but the compatibility with the hypothesis of a stationary property requires investigation. If we considered that the principle of hydrodynamic continuity were valid over large regions and with perfect accuracy then it would follow that the mean density of matter was decreasing, and this would contradict the perfect cosmological principle. It is clear that an expanding universe can only be stationary if matter is continuously created within it. The required rate of creation, which follows simply from the mean density and the rate of expansion, can be estimated as at most one particle of proton mass per litre per 10^9 years. In interpreting the universe as stationary we have to assume that such a process of creation is operative; we have to infringe the principle of hydrodynamic continuity. But this principle is not capable of experimental verification to such a precision, and this infringement does not constitute a contradiction with observational evidence. It is true that hydrodynamic continuity has been regarded as an unqualified truth and not as an approximation to physical laws, but this was merely a bold simplifying extrapolation from evidence. Hydrodynamic continuity is no doubt approximately true but this does not compel us to assume that it holds without any deviation whatever. In the conflict with another principle which is much more far-reaching and capable of making many more statements about the nature of the universe and the applicability of physical laws, there is no reason for upholding the principle of continuity to an indefinite accuracy, far beyond experimental evidence.

We have referred to the principle infringed by the creation process as the principle of hydrodynamic continuity; as we shall see later it is by no means clear whether we can regard any principle of conservation of energy as infringed.

2.4. Much of the structure of the universe is closely defined by the perfect cosmological principle and the knowledge of the velocity-distance law of expansion. As we shall see later such a universe must be described by a de Sitter metric. The universe is therefore spatially as well as temporarily infinite, but the effect of distant matter on an observer tends to zero as its velocity of recession approaches the velocity of light.

The motion of matter in such a universe follows a particular pattern; though in the actual case there are fluctuations, there is a preferred motion defined at each place. Any observer equipped with a sufficiently good telescope can make observations of the velocity-distance law in different directions, and from such observations he can deduce any deviation that he may have from this preferred motion in his locality. If he follows this motion without any deviation then he will see symmetry around him; although all the objects which he observes may have some deviations from the preferred motions in their localities, a statistical averaging method can serve to specify the preferred motion at the observer exactly. (It is not necessary to specify the averaging method at this point beyond stating that it must give sufficient weight to distant matter.) At any rate there is not much latitude. Such a preferred motion could be established by an observer at any place, and the resulting pattern forms a vector field in a four dimensional space-

time. Much physical significance must be attached to this vector-field, for the motion of all matter appears to be closely dictated by it. What deviations there are, are statistically distributed around the preferred motion, and these deviations are quite small compared with the velocity of light.

An interesting problem connected with the preferred direction is that of the existence of a "cosmic time". In the cosmologies of general relativity the existence of a cosmic time is deduced from Weyl's postulate (**2**). As observations show, there are irregularities and deviations from uniformity. These are only small, but if they are such that there are no surfaces orthogonal to the world lines of all the nebulae, then there is no cosmic time. The existence of a cosmic time in the presence of these irregularities has not been proved and seems improbable. The ambiguities in cosmic time arising from such a breakdown are only small in any region, but over the universe as a whole they completely undermine the logical structure on which a homogeneous but not stationary theory rests. Our position is very different. Owing to the stationary character of the universe, homogeneity can be defined in the absence of cosmic time. Whether or not there are surfaces orthogonal to the preferred direction is therefore of no importance for our theory. On general grounds one would not expect these surfaces to exist, so that an unambiguous definition of cosmic time is impossible. In the simplified smoothed-out model it will of course exist, but probably not in the actual universe with its irregularities.

The stationary property implies that the factor of the velocity-distance law must stay constant. This would not be fulfilled if matter moved such that the velocity between two particular masses remained constant; the velocity has to increase as the distance increases. Accordingly, there must be relative acceleration between matter.

We can now examine the requirements which the perfect cosmological principle places on the evolution of stars and galaxies. The mean ratio of condensed to uncondensed matter has to stay constant, and for this reason new galaxies have to be formed as older ones move away from each other. We take the perfect cosmological principle to imply that no feature of the universe is subject to any consistent change, and no observer hence capable of any unique definition of a universal time. This will be satisfied only if the ages of galaxies in any sufficiently large volume follow a certain statistical distribution. With such a distribution it is of course possible for any particular volume to contain a very old galaxy; but very old galaxies are thinly scattered, as they have had time to move far apart in the general expansion. The characteristic time in this age distribution must of course be related to the characteristic time of expansion of the universe. In opposition to most other theories we should hence expect to find much diversity in the appearance of galaxies, as they will be of greatly different ages; and much diversity is in fact observed. Furthermore the age distribution of galaxies in any volume will be independent of the time of observation, and it will hence be the same for distant galaxies as for near ones, although in the former case the light has taken long to reach us. This is a property of great importance, for it is in practice the only basis on which we can construct any interpretation of the observations of distant galaxies.

2.5. The next point to discuss is the thermodynamic state of the universe. It is known that this state is very far from any equilibrium. Very much more energy is in the form of matter than in radiation; and very much more energy is radiated away than is absorbed by matter.

The disequilibrium is very great indeed, and we have grown so used to it that we take its existence to be self-evident. Such statements as "A hydrogen atom in an excited state returns quickly to its ground state, since this has less energy", do not indicate any lack of faith in the principle of conservation of energy but merely an awareness of the paucity of radiation in the universe which makes the reverse process so very unlikely.

This disequilibrium is such a marked and characteristic feature of our universe that its explanation must be one of the main tasks of cosmology. The connection between this phenomenon and cosmology was noticed by Olbers (**3**), who pointed out that in an infinite homogeneous static Newtonian universe the mean radiation density would be as high as it is on the surface of a star! In the present case, however, this disequilibrium is the direct consequence of the motion of expansion. Any photon emitted by a star has a very large probable free path, owing to the tenuous distribution of opaque matter in the universe. The most likely career for such a photon is that it will travel on into regions where the local motion of matter is very different from the motion of the star which caused its emission. When such a photon is finally absorbed, then it will, as seen by an observer on the absorbing matter, be subject to a Doppler-shift which reduces its frequency by a large factor. The thermal energy it will supply to the intercepting matter will hence be only a small fraction of the energy it removed from the emitting star. It is not necessary to consider here whether the difference in energy can be regarded as being supplied to the expansion process of the universe; but it is clear that the universe provides in this way a sink for radiative energy, and that this sink is in fact available to the majority of photons produced at the surface of stars. The process of getting rid of both matter and radiation from any fixed volume is by pushing both across the surface bounding this volume; and both are replenished from within.

It is of course important for any theory of cosmology that it should not predict thermodynamic equilibrium, and this is possible only by supposing either that insufficient time has been available since the commencement of radiation by stars, or, and this is the view we take, by suggesting that the bulk of the energy radiated does not again become available as heat.

A thermodynamic disequilibrium is a requirement for any local thermodynamical evolution, and in this sense it is necessary in order that any physical significance should be attached to the passage of time. In a stationary universe this condition can be fulfilled only if there is a motion of expansion, and if the mean free path of a photon is at least comparable with the dimension defined by the velocity-distance law and the velocity of light. (The radius of a de Sitter universe.) Expressing the same consideration differently: the perfect cosmological principle and the knowledge of the thermodynamic disequilibrium require that the universe should be expanding. The experimental evidence of the recession of nebulae is hence a good indication that this specific requirement of the perfect cosmological principle is in fact satisfied. The narrow cosmological principle would have resulted in no such requirement, for there the disequilibrium would result simply if the luminosity of all stars was a suitable function of universal time.

Let us now consider the question of the overall conservation of mass (energy) in such a universe. An observer attempting to estimate the quantity of matter in the universe will find that his result depends on the threshold intensities and wavelengths of his apparatus. Keeping these fixed, he will always observe a finite amount of matter and this will be constant in time. At a great range matter is

drifting into an unobservable state by approaching the velocity of light, and without a process of creation this would not allow any principle of conservation of energy to be applied to the sum total of all observable matter.

A further point to be mentioned in relation to the stationary property of the universe is the coincidence of numbers pointed out by Eddington (**4**). Two non-dimensional numbers which can be constructed from observation are both found to be of the order 10^{39}; one is the ratio of the characteristic length defining the expansion of the universe to the classical radius of the electron (e^2/mc^2), and the other is the ratio of the electric to the gravitational force between a proton and an electron. The coincidence is striking, and though we have no causal connection to offer we feel that one exists. This and some other coincidences cannot be ignored by a system of cosmology to the extent of suggesting that they occur only in a fleeting phase. It is satisfactory that according to the theory presented here such coincidences are permanent.

3. *The Observational Tests*

3.1. Our observations of the universe depend on the interactions between the rest of the universe and ourselves. These interactions are of two kinds: dynamical and electromagnetic. Dynamically, the rest of the universe affects us primarily through the long-range force of inertia, i. e. it defines inertial frames of reference in our neighbourhood. This dynamical interaction is strong and arises almost certainly from the whole of the observable universe, but it is ill-understood and is discussed in other parts of this paper. Therefore, in this section we shall confine ourselves to a discussion of the electromagnetic interactions, which are more easily interpreted.

Light reaches us in observable intensities from large numbers of extra-galactic nebulae, and a considerable amount of information has been gathered. In particular the distribution over the sky and the number of nebulae of given apparent magnitude is quite well known, and the radial velocities have been determined (from the Doppler shift) for many nebulae.

In trying to interpret these data, we must first decide whether the observations extend over a sufficiently large region to have any bearing on cosmology, or whether they must be taken to refer to a purely local district.

There are two strong arguments (isotropy and homogeneity of nebular distribution and magnitude of the radial velocities), for assuming that the observations cover a fair sample of the universe and there is no evidence against it. In common with most writers on the subject we assume therefore that the observations refer to a sufficiently large region to be of relevance to cosmology.

3.2. We must now see whether our theory, at this stage of its development, can make any statements about the aspect of the universe which can be compared with observation. The mathematical apparatus we can use is strictly limited, since the field equations of the general theory of relativity with their insistence on conservation of mass are clearly not acceptable to us. In fact we must rely largely on kinematical arguments and on the generally accepted results of laboratory physics dealing with the propagation of light (special theory of relativity). The power of these kinematical arguments was demonstrated by the pioneer work of Milne and by the papers of Walker and Robertson.

A simplified model of the universe can be obtained by assuming the conditions of homogeneity and isotropy to be absolutely (instead of only on a large scale)

satisfied. There seems to be good reason to believe that the features amenable to long distance observation are nearly the same in this simplified model as in the actual universe with all its irregularities. In common with most authors we shall assume that this representation of the universe is sufficient for comparison with the observations of extragalactic nebulae.

The mathematical properties of such a model have been studied in great detail by Robertson (**5**) and by Walker (**6**). These authors adopt a narrower cosmological principle and they do not assume a stationary property. They do, however, assume that the particle paths do not intersect except that possibly they all intersect at one singular point in the past (origin of universe). Although we do not wish to make this assumption in general it is clear that in the simple model we are now discussing the assumption will be satisfied. For if there were an intersection at a point P, then the isotropy condition would show that this point had to be a focus, and the homogeneity condition would then lead us to a universe in which every point is a focus. Such a universe is evidently so different from the one we know that we need not hesitate to reject this possibility.

Accordingly, the work of Robertson and Walker is applicable to our theory, except that we have to specialize their results by the stationary condition. In this way the following theorem can be established * :

In a non-static stationary universe the light-paths are the null geodesics of the metric

$$ds^2 = dt^2 - (dx^2 + dy^2 + dz^2) \exp(2t/T),$$

where the (x, y, z) coordinates of any particle partaking of the general motion of the universe are constant and t measures the proper time of any observer travelling with such a particle. T is a universal constant and the velocity of light has been put equal to unity. This metric is very well known, having been discovered by de Sitter (**7**) in 1917. It is the simplest of all metrics of the Robertson type, but could not be used in general relativity, since the field equations imply that it describes empty space only.

The geometrical properties of the de Sitter universe are very simple. If we change from Cartesians (x, y, z) to spherical polars (r, θ, ϕ) in the usual way, then an observer at the origin would find that the spectral lines of a source of light of luminosity L at (r, θ, ϕ) received by him at time $t = 0$ would show a red shift

$$1 + \delta = \frac{\lambda_{\text{received}}}{\lambda_{\text{emitted}}} = \exp(-t_{em}/T) = 1 + r/T, \tag{1}$$

while the intensity of light received would be

$$\frac{L}{4\pi r^2} \exp\left(\frac{2t_{em}}{T}\right) = \frac{L}{4\pi r^2 (1 + r/T)^2}. \tag{2}$$

These results are well known (**8**) and apply equally in general relativity and in our theory. An important difference occurs, however, when we consider the third observable quantity, the number of nebulae in a given magnitude class. In general relativity it is assumed that the density of nebulae was greater in the past than it is now, owing to the expansion of the universe. The theory here presented assumes that the density was the same in the past as it is now. Accordingly, if n

* Another way of proving the theorem is to assume the existence of a Riemannian metric and then to apply the conditions demanded by our postulates.

is the number of nebulae per unit volume, the number of nebulae between r and $r+dr$ is

$$4\pi r^2 n dr \exp(3t_{em}/T) = 4\pi r^2 n dr(1+r/T)^{-3}. \quad (3)$$

The comparison with observation consists in checking that the interdependence of the functions (1), (2), (3) is the same in observation as in theory. This process is familiar from relativistic cosmology, but there is another difference not yet referred to. In our theory the average nebular luminosity L must be constant by virtue of our fundamental assumption. This is not so in any other cosmology, where the distant nebulae, from which light reaches us only after a long time, must be assumed to have been in an earlier stage of development when the light observed was emitted. Owing to the likely variation of nebular luminosities the measurement of distances becomes very inaccurate just at the ranges where space curvature begins to be important. The theory here presented is therefore the only one which can be compared with observation without having to make any assumptions about nebular evolution.

3.3. In all comparisons between theory and observation difficulties arise owing to the fact that we always measure photographic luminosities, while the theory is concerned with bolometric luminosities. The difference between the two quantities depends on the nebular spectrum and is a function of the red shift. This so-called nebular K term is not very precisely known. A recalculation was carried out giving results in good agreement with (though slightly lower than) Hubble's results (**9**).

The velocity shift can only be measured for fairly bright nebulae ($m<17{\cdot}5$) and is found to be

$$\log_{10}\delta = 0{\cdot}2m - X, \quad (4)$$

where m is the magnitude (defined as const. $-2{\cdot}5\log_{10}$ (apparent luminosity)). X is a constant which Hubble and Humason (**10**) gave as 4·967 in 1931, while Hubble (**11**) deduced the value 4·707 in 1935. ten Bruggencate (**12**) considers 4·907 to be the most probable value. The difficulty in the determination of X is due to the paucity of measurements in the middle region ($14<m<17{\cdot}5$) and the obscuring effects of the peculiar motions in the near region ($m<14$).

The surveys of the number of nebulae brighter than a given limiting magnitude constitute the principal information for the fainter nebulae. The following table gives the results of the surveys (**13**). $N(m)$ is the number brighter than magnitude m over the whole sky, extrapolated from the samples taken.

m	12·78	18·2	18·47	19·0	19·4	20·0	21·03
$\log_{10} N(m)$	3·233	6·328	6·502	6·774	6·955	7·301	7·777
$0{\cdot}6m-\log_{10} N(m)$—const.	0	0·157	0·145	0·189	0·250	0·264	0·406

The theoretical relations derived from (2), (3) and (4) are

$$m = K + 5X + 5\log_{10}\delta(1+\delta), \quad (5)$$

$$N = A\left[\log_e(1+\delta) - \frac{\delta}{1+\delta} + \frac{\delta^2}{2(1+)\delta^2}\right], \quad (6)$$

where K is the term previously referred to and A is a constant. The expression $0{\cdot}6m-\log_{10} N(m)$ is the most sensitive indication of the departures from apparent uniformity (which are due to the recessional velocities).

The theoretical curves of $0{\cdot}6m - \log_{10} N(m)$ against m are plotted for various values of X and the observational results are indicated by crosses in Fig. 1. It will be seen that the agreement between theory and observation is extremely good if the larger value of X is adopted, but is only moderately good for the lower value of X.

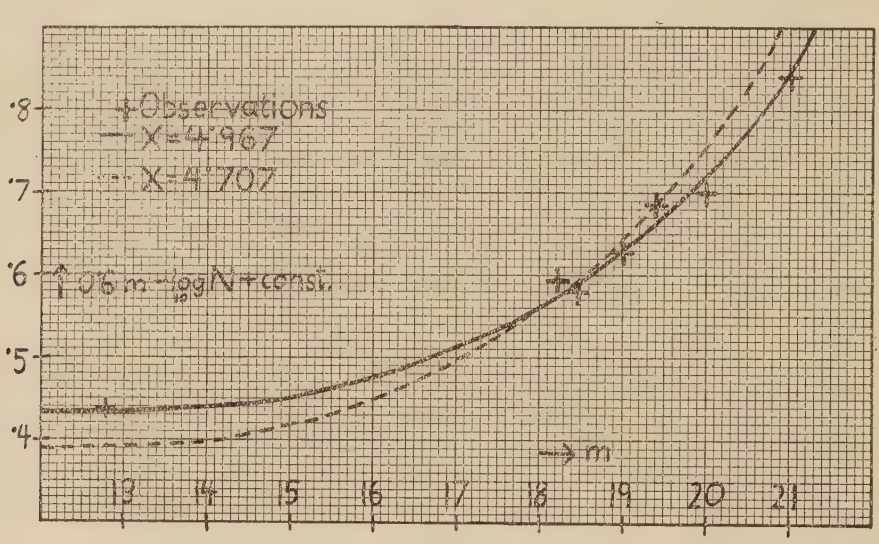


FIG. 1.

Before proceeding to an examination of the implications of these data we must contrast the purpose of these comparisons in our theory and in general relativity. In general relativity a very wide range of models is available and the comparisons merely attempt to find which of these models fits the facts best. The number of free parameters is so much larger than the number of observational points that a fit certainly exists and not even all the parameters can be fixed.

In our theory, on the other hand, the parameters are determined by observations of near nebulae. The counts of distant nebulae serve merely as a check on the theoretical forecast.

It is clear from the figure that this agreement is as good as can be expected from the uncertainty of the data. The possibility of systematic errors in the determination of the limiting magnitudes and the uncertainty in the K term may easily account for the slight systematic divergence which a confirmation of the lower X value could otherwise suggest. A confirmation of the higher X value on the other hand gives in any case perfect agreement.

We see therefore that our theory agrees with observation at least as well as the relativistic models do when their numerous free parameters have been adjusted. Hubble's original deduction that his data suggested, on a relativistic interpretation, an improbably small universe is qualitatively in good agreement with our theory : the paucity of faint nebulae, which in relativity must be ascribed to positive space curvature, is on the basis of our theory simply explained as being due to a density of nebulae which is constant in time, instead of decreasing as hitherto assumed.

3.4. A difficulty which occurs in other cosmological theories is absent in ours : the problem of the time-scale. It will be useful to discuss at some length the difficulties associated with the time-scale in other theories.

The reciprocal of Hubble's constant, which is a time (T), defined by the observations of the velocity-distance law, is between 1800 and 2000 million years. It is clear that this time must be of fundamental importance in any theory.

Terrestrial observations of radioactive decay in rocks indicate an age of these rocks of at least 2×10^9 years, and some samples seem to imply an age of more than 3×10^9 years. Astrophysical considerations tend to indicate an age of the stars in our galaxy of about 5×10^9 years.

How do these results affect cosmology? If the continuous creation of matter is denied then the following two possibilities exist :

(i) The rate of expansion of the universe in the past was the same as now or greater. Then the origin of the universe must have been catastrophic and took place less than 2×10^9 years ago. This clearly contradicts the terrestrial and astrophysical evidence referred to above. A distinction between a "dynamical" and a "cosmic" time, as suggested by Milne, removes only part of the difficulty. It makes some of the astrophysical evidence inadmissible but, since atoms and nuclei must be supposed to follow the cosmic time, the terrestrial evidence remains.

(ii) The rate of expansion was slower in the past than it is now. Then some disruptive force must have been responsible for the acceleration. Such a force is described by the hypothetical and much debated cosmological term in general relativity. A general criticism of any model based on such a disruptive force can be made, but it will be useful to discuss a special example, viz. the famous model due to Lemaître (**14**). In this model the universe is supposed to have been for an arbitrary long time in the unstable Einstein state in which λ repulsion exactly balanced gravitational attraction. The formation of condensations led to an expansion proceeding at an exponentially increasing speed.

Much more attention has been paid to the attractive formal features of this model than to the severe objections to which it is open on physical grounds.

Although its past is mathematically infinite it must be taken to be entirely featureless and void of any physical development. For any physical process such as the formation of a condensation would have upset its equilibrium. In fact in its "pre-historical" state (before the first condensation) circumstances were such that the formation of a condensation was only just possible, though very unlikely, since otherwise it would have occurred earlier. However, as soon as the physical history of the model started expansion began and so the mean density decreased. Very early on in the expansion this decrease would have been sufficient to make the previously improbable formation of condensations quite impossible. Therefore the period of physical evolution of this model is brief in spite of its infinitely long mathematical existence.

The unsatisfactory features of this model are shared by the other accelerating models of general relativity. The theory presented here differs from all these models very profoundly. Although the universe is accelerating in the sense that the speed of each particular nebula is increasing, the mean density remains constant. Therefore conditions for the condensation of new nebulae are always the same and may be taken to be always moderately favourable. The probability of the formation of condensations is presumably determined by the important non-dimensional quantity $\gamma\rho T^2$. In our theory this is a constant, since all three factors are constant.

(In general relativity it is variable, but in kinematical relativity it is constant, although all the factors vary.) Observations indicate that the value of this constant is between 10^{-3} and 1, probably nearer 1.

The time T is an important constant of nature in our theory, but it has little to do with the age of our own galaxy. The development of the galaxies and their motions are sketched in Fig. 2.

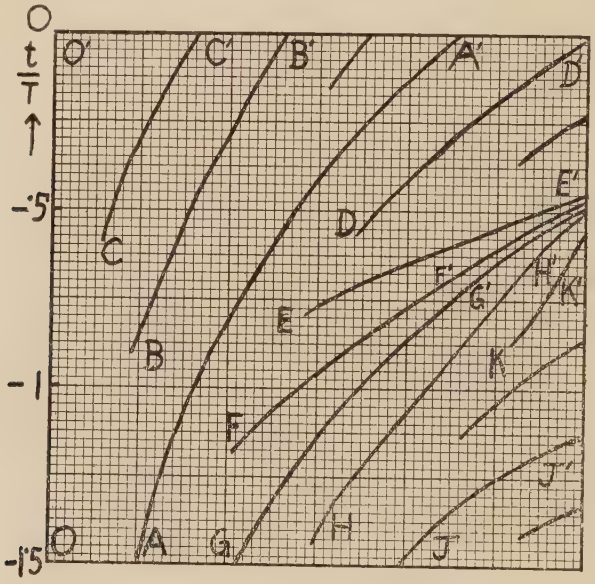


Fig. 2.

The time-axis is drawn vertically, T being taken as unit of time. The abscissa represents distance suitably defined. The observer is taken to be on the old galaxy OO′, which follows the time-axis.

At $t = -T$ the galaxy represented by the world line AA′ was O's nearest neighbour. By $t = 0$ it has drifted far away, but new galaxies (represented by BB′, CC′, etc.) have sprung up in the meantime. Occasionally galaxies capture each other gravitationally (FF′, GG′) and then, owing to their combined mass, are likely to capture many more (HH′, KK′). All the time, however, this cluster is moving away from O.

The ages of the nebulae follow therefore a merely statistical law and there is no reason to suppose that a particular nebula (such as our Milky Way) is of some age rather than another. In our theory there is therefore no difficulty whatever in taking the age of our galaxy to be anything indicated by local observations (such as say $5–8 \times 10^9$ years), although T is much shorter.

The statistics of such a system are difficult, but very tentatively the course of evolution may be pictured as going from the single field nebula to systems such as our own with a dozen or two members and ending up as the great clusters, with a population of hundreds. The older the system the more likely it is to be far away from any other old galaxy, so that at any time only a limited number of great clusters should be within regions which can be explored with a telescope of given size.

Although this view of nebular evolution does not form an integral part of the theory, we should mention why we do not follow the view that clusters are evaporating rather than condensing. The time constant for gravitational capture is $(\gamma M/d^3)^{-\frac{1}{2}}$, where M is the mass of a nebula and d the internebular distance.* This time is therefore only about 10^{10} years, compared with the period of 10^{12} years required for evaporation by dynamical friction (**15**). Furthermore, the presence of even a very tenuous intergalactic gas would not allow nebulae to acquire the "accidentally" high velocities required for evaporation. Therefore we expect that in general the tendency for evaporation will be negligible compared with the gravitational clustering tendency.

If we assign to our galaxy, on the astrophysical evidence, an age of between 5×10^9 and 10^{10} years, it follows that there should be a moderate degree of clustering such as is actually observed.

4. *The Physics of Creation*

4.1. We shall now attempt to consider some of the details of the physical process of creation. The discussion must necessarily be somewhat speculative but some deductions can be made which limit the range of the possibilities. The average rate of creation is determined by the rate of expansion and the density of the universe (**16**) and is approximately 10^{-43} g. per sec. per cm.^{3}. We now must discuss how this creation rate is distributed in space. Will it be roughly the same everywhere or will it be very much greater in some regions than in others? In particular will it be as concentrated as the existing matter is (in which case it would be more appropriate to speak of creation per unit mass per unit time)?

The answer to this question is intimately connected with the theory of evolution. There are, roughly speaking, at present two schools of thought in this matter :

(i) One body of astronomers is of the opinion that the main line of development is centrifugal. For example, they consider that nebulae evaporate from clusters and so form the general field, that far more matter is ejected by stars than is accreted by them, so that interstellar space is populated primarily by matter which once was stellar, etc.

(ii) Another body of astronomers considers that centripetal development is predominant. In their view, accretion is far more important for stellar evolution than the ejection of matter from stars; they consider the gradual condensation of stars in the interstellar material (which is taken to constitute the bulk of the mass of a nebula) to be the main process of nebular evolution, etc.

Now it is clear from the discussion in Section 1, that in order to ensure the stationary aspect of the universe, the new material must be created in the form from which evolution starts. For, if evolution turns matter from state A into state B, then the proportion of matter in state A will be diminished unless the newly created matter is in state A. Therefore the created matter must start at the beginning of the evolutionary chain to replenish the reservoir that would otherwise dry up. Accordingly if we assume that creation is by unit mass, then we have to assume that the centrifugal line of thought is correct over an even wider field than its adherents claim, while if we assume the rate of creation per unit volume to be about the same everywhere, then we have to follow the centripetal school.

* Two equal nebulae of mass M originally at rest relatively to each other at distance d would collide after a time $t=\frac{1}{4}\pi(\gamma M/d^3)^{-\frac{1}{2}}$

In our opinion the first view is untenable. For if matter were created mainly in the stellar interior it would have to boil off or be ejected in explosions to a distance and at a rate sufficient to supply intergalactic space with enough material for the required rate of formation of new galaxies. This implies that on an average each star loses its entire mass, by such a process of evaporation or explosion, in $\frac{1}{3}T\log_e 2 = 4 \times 10^8$ years! The known rate of occurrence of novae and supernovae shows that these phenomena can only make an insignificant contribution to this. This calculation given here does not allow for the necessary recondensation of the ejected material in its original galaxy, so that the period would be even shorter. This conclusion seems to be so absurd that we take it to prove that the rate of creation per unit volume per unit time is roughly constant and that evolution proceeds mainly in the centripetal direction.* It must be clearly understood that some variation of the rate of creation is still permitted, it is only necessary that far more matter is created in space than in the stars.

Accordingly we assume that the rate of creation is everywhere about 10^{-43} g. per sec. per cm.^{3}, although variations by a factor of, say 100, are quite permissible. This rate corresponds to one new atom of hydrogen per cubic metre per 3×10^5 years. Accordingly we cannot expect the process to be directly observable.

4.2. We have seen that the process of creation will take place everywhere and that most of the matter will be created in the vast intergalactic spaces. The next question to engage our attention will be the initial motion of the created matter.

It is evident that this velocity vector associated with creation cannot be relativistically invariant, nor can invariance be introduced by assuming a random distribution of velocities, since the only invariant distribution diverges.

It follows, therefore, that there will be a preferred time-like direction at each point of space-time, and that the initial motion of newly created matter will be either along this direction (in the 4-dimensional sense) or will be distributed about it in a non-invariant way. The question of this distribution will be discussed later; for the moment it will be assumed that the matter is actually created along the preferred direction.

The adoption of a non-invariant law of creation may at first sight seem surprising; but are we entitled to demand invariance of creation when the resulting movement of matter is well known not to be invariant? Even in the cosmologies of general relativity the existence of a preferred direction is explicitly recognized in Weyl's postulate. General relativity demands that the laws of cosmology should be invariant while admitting that the one and only application is not invariant. We can see no reason why the laws of nature determining the structure of our universe should be invariant, although the universe is unique and does not bear an invariant aspect.

The identification of the preferred direction defined by the creation of matter with the preferred direction defined by the motion of matter (as discussed in 2.4) is compelling. It is most satisfactory that our theory can connect the direction defined everywhere by the motion of matter with another phenomenon, while in other theories no such connection exists.

It should be realized that the existence of this preferred state of motion not only conflicts with general relativity, but also with restricted (and even with Galilean) relativity. However, this lack of invariance is nothing new. As we have seen in

* This conclusion has already been anticipated in the discussion on clusters in 3.4.

Section 2 it is an immediate consequence of the law of the recession of nebulae.

4.3. We must now consider the problem of the probability distribution of the initial velocity about the preferred direction. If the spread of velocities were great, then this would lead either to large amounts of radiation from intergalactic space or to high temperatures of intergalactic matter.

The amount of radiation observed and the postulated intergalactic condensation limit the velocity spread to at most 10^7 cm./sec., corresponding to a temperature of at most 50,000 deg.

There do not seem to be any arguments setting a lower limit to the temperature and the arguments limiting the temperature from above all depend on the absence of any effects of the temperature. If one wishes to have a definite and simple picture of the process, one can therefore assume that matter is created without random velocities (i. e. at zero temperature), although this is not a necessary conclusion.

4.4. The next question we must discuss is the electric charge of the created matter. Now it is clear that an average excess charge of either sign would lead to forces of repulsion tending to remove the excess charge to infinity. The speed with which the excess charge is repelled to infinity depends on the ratio of excess charge to mass. It can be shown (**17**) that if this ratio is greater than about $10^{-19} \times$ electron charge/electron mass, then this repulsion of excess charge would take place at a speed greater than the speed of expansion of the material universe. But such a rapidly "expanding universe of excess charge" cannot be present together with a slowly expanding "universe of matter" without violating the principle of homogeneity. Accordingly we can conclude that if there is any average charge excess of the created matter, the average charge-mass ratio must not exceed the very low limit mentioned above.

These results give us some indication of what the created matter may be. If we reject as far-fetched the possibility of the creation of dust, etc., and concentrate on the simplest building material, then we are left with the following possibilities :

(i) Protons and electrons, created separately but at identical (or very nearly identical) rates. The radiation arising from collisions or captures would be just within the limits laid down in 4.3.

(ii) Neutrons. The energy released by the disintegraton of the neutrons would (within a few hours of creation) lead to velocities and radiation somewhat in excess of the limits mentioned. Nevertheless, this possibility is so attractive that it should not be ruled out without further examination.

(iii) Hydrogen atoms.

The last possibility seems to agree best with the view adopted in 4.3, although (i) cannot be excluded. Since (i) would, but for the necessarily small amount of radiation, also lead to the formation of hydrogen in the ground state, we can for simplicity and definiteness adopt (iii).

In agreement with the widely held view that the building-up of the elements is a complex stellar process and that the traces of elements other than hydrogen in interstellar space are due to explosion and ejection of matter from stars, we have assumed here that all matter starts its career as hydrogen. This seems to us to be very plausible, but is a separate subject not directly connected with the creation theory.

4.5. We finally have to consider where in space-time matter is created and it seems to be most reasonable to assume that it is created in a random manner.

According to this view the probability of creation taking place in any particular four-dimensional element of volume (spatial volume element × element of time) is simply proportional to its (four-dimensional) volume, the factor of proportionality being a function of position.

By our argument in 4.1. this factor cannot vary very much from point to point.

Accordingly randomness would be introduced by the random origin of particles, but not necessarily by any initial random motion.

5. *The Function of a Field Theory*

5.1. We have now completed that part of the development of our theory which deals with the motion of the universe as a whole and with the physical character of the new concepts introduced. The next step must be a more detailed formulation of the theory in terms of which it is possible to discuss the motion of individual particles not necessarily following the general expansion of the universe. Two methods of attack on this problem are possible. One can either follow the methods of kinematical relativity in order to examine the consequences of the perfect cosmological principle, or one can try to build up a field theory. The first possibility has been well explored for a less stringent cosmological principle by Milne, Walker and their school. Our perfect cosmological principle should make progress easier and additional assumptions less necessary.

It is, however, known from kinematical relativity that the application of this technique is by no means easy unless the cosmological principle is assumed to apply to every detail of world structure. This is of course not true; the cosmological principle is a purely statistical law in the universe, applying only to large-scale averages. The need to make this additional and unwarranted assumption makes the interpretation of the theory difficult and its applicability somewhat doubtful.

We must stress here that in our view the difference between statistical and complete homogeneity imposes a serious requirement on any theory of cosmology. While it is quite legitimate to consider first a completely homogeneous universe for reasons of mathematical convenience, great care must be taken to see that the introduction of fluctuations affects only the mathematical detail and not the logical structure of the theory. This requirement is not fulfilled by a number of the models of general relativity.

The development of a field theory is therefore of particular importance. We wish to stress that in our view general relativity cannot be satisfactorily adapted to our theory, in spite of its great attractiveness. Though general relativity is presumably correct in stating that the light cone, together with space and time measurements, defines a tensor field throughout space-time, we reject entirely the claim that this is the only fundamental covariant field universally defined. An equally fundamental vector field is defined by the state of motion in which an observer must be in order to see isotropy in the universe around him. This universally defined vector field exists in every cosmology, but it is of particular importance in the theory presented here. This was discussed in detail in 3.1. In our opinion, this field should play as important a part in the theory as the tensor field.

As a consequence of general relativity's neglect of this vector field, it is far too wide a theory, i. e. it covers a far greater range of possibilities than actually exist. This is frankly recognized in relativistic cosmology and Weyl's postulate is used to narrow it down sufficiently. To us this narrowing-down of the theory in its final

form seems to be utterly unsatisfactory, these restrictions should enter the theory at the beginning and not at the end.

Another fundamental objection to general relativity is that it does not satisfy Mach's principle completely. Finally, the strict conservation of mass implied by the field equations conflicts with the very basis of our theory. This has already been pointed out in Section 2.

5.2. The authors wish to express their gratitude to Mr F. Hoyle for many helpful discussions they had with him on the question of a field theory formulation of the perfect cosmological principle. His early appreciation of the significance of the continuous creation of matter has been a great encouragement to us. He is publishing (**18**) a most interesting adaptation of the field equations of general relativity which does not conflict with the idea of creation and has the de Sitter universe with constant density and zero pressure as a solution.

This formulation has a number of attractive features, chief amongst which is the smallness of the change in the field equations. This permits the use of the whole apparatus of general relativity. Nevertheless, we feel that this formulation is unsatisfactory and unacceptable.

A field theory is only required to deal with deviations from perfect uniformity, since, as was shown in this paper, the case of uniformity can be dealt with easily without field theory. But the insertion of non-uniformities raises very serious objections to Hoyle's theory. For every concentration of mass will deflect the pencil of geodesics corresponding, in his theory, to the particles before they are created. This will create multiplicities in the C field, which a very long time after the occurrence of his event O, will be of very high order. At first sight one would expect this to imply an impossibly high temperature of the created particles, far in excess of the limits laid down in 4.3. Mr Hoyle has however very kindly informed us that this is not the case. He has found that the effect of his C field is to deflect the orbits of particles so that they are no longer geodesics. The deflection is always such as to increase uniformity, i.e. relative velocities are diminished. Furthermore, massive bodies are far less strongly affected than a tenuous gas. Therefore this "C force" has little effect on stars, but it strongly deflects newly created particles which form a gas of very low density, so strongly in fact that by the time the first collisions are likely to occur the kinetic temperature is already low.

While this argument shows that the theory survives the temperature test, it largely destroys at the same time its attractiveness. The geodesic postulate plays a great part in the development of general relativity. In Hoyle's theory it is dropped in such a way as to infringe even Galileo's principle that all bodies fall equally fast. However, the infringement is far too small to be detected by practical astronomy. The direct physical reality given to space itself by the C field is therefore very great indeed. Hence the C field exerts forces on material bodies. While this is necessary to give reality to the vector field discussed in 2.4, it seems unsatisfactory to us that the C field is only partly determined by the existing motion of matter and is largely determined by Hoyle's event O in the infinite past.

Some of the objections made here against relativity can be made equally well against Hoyle's theory. In particular Mach's principle is satisfied no more and no less than in general relativity.

In view of these objections we have no hesitation in rejecting Hoyle's theory, although it is the first and at the moment only field theory formulation of the hypothesis of continuous creation of matter.

5.3. We hope to present soon in another paper a field theory formulation free from these objections. It may however be advantageous to state here already why, in our view, the principles laid down in the present paper should enable one to reformulate the entire problem of inertia and the associated acceleration of the galaxies. The need to explain the self-perpetuating character of the universe rather than the "force" responsible for its expansion will presumably dispense with the necessity for such an artificial concept as the λ force of general relativity. Furthermore, in all field theories hitherto presented, the interaction between a particle and the bulk of the universe had to be assumed to be weak and indirect, because of the need to have permanent intrinsic laws of nature in a supposedly changing universe. There is no such restriction in our case. The stationary character of the universe permits us to assume very strong interactions and yet to have permanent laws of nature. Only in this way can Mach's principle be truly satisfied and only such a theory can establish satisfactory connections between cosmology and laboratory physics.

Trinity College,
Cambridge:
1948 *July* 13.

References

(1) Mach, E., *The Science of Mechanics*, London, 1893, pp. 229–238.
(2) Weyl, H., *Phys. Z.*, **24**, 230–232, 1923; *Phil. Mag.*, Series 7, **9**, 936–943, 1930.
(3) Olbers, *Bode's Jahrbuch*, 10, 1826.
(4) Eddington, A. S., *Proc. Roy. Soc.* A, **133**, 605, 1931; *M.N.*, **92**, 3, 1932.
(5) Robertson, H. P., *Ap. J.*, **82**, 284, 1935.
(6) Walker, A. G., *Proc. Lond. Math, Soc.*, Series 2, **42**, 90, 1936.
(7) de Sitter, W., *Proc. Akad. Wetensch. Amst.*, **19**, 1217, 1917.
(8) Robertson, H. P., *Rev. Mod. Phys.*, **5**, 71, 1933.
(9) Hubble, E., *Ap. J.*, **84**, 537, 1936.
(10) Hubble, E. and Humason, M. L., *Ap. J.*, **74**, 43, 1931.
(11) Hubble, E., *Ap. J.*, **84**, 270, 1936.
(12) ten Bruggencate, P., *Naturwiss*, **25**, 561, 1937.
(13) Hubble, E., *Ap. J.*, **84**, 517, 1936. Shapley, H., *Proc. Nat. Acad. Sci.*, **24**, 148, 1938. Shapley, H. and Ames, A., *Ann. Harv. Coll. Obs.*, **88**, No. 2, 1932. The treatment given here is adapted from McVittie, G. C., *Proc. Phys. Soc.*, **51**, 529, 1939.
(14) Lemaître, G., *M.N.*, **91**, 483, 1931.
(15) Tuberg, M., *Ap. J.*, **98**, 501, 1943.
(16) *Cf.* Hubble, E., *The Realm of the Nebulae*, Oxford, 1936, p. 189.
(17) Unpublished theorem of Newtonian Cosmology.
(18) Hoyle, F., *M.N.*, **108**, No. 5, 1948 (in press).

THE SPECTRUM OF ρ CASSIOPEIAE IN 1939

W. S. Tai and A. D. Thackeray

(Received 1948 April 19)

Summary

Two hundred and sixty-four lines between 4226 and 6374 A. in spectra of ρ Cassiopeiae obtained with two-prism and four-prism spectrographs at Cambridge in 1939 have been measured. Identifications are presented in a summarized form and a comparison made with the supergiant α Persei. Rare-earths, *Sc*II, *Zr*II and *Y*II are strong in ρ Cass, *H* and *Fe* are weak. Unblended lines yield a radial velocity of $-46{\cdot}9$ km./sec.

ρ Cassiopeiae is an irregular variable star whose amplitude of variation has been given by various observers (**1**) as ranging from $0^m{\cdot}15$ (in the years 1927–1931) to $0^m{\cdot}7$ (in the years 1916–1923), and whose period has been given as ranging from 200 to 1100 days. Estimates of its spectral type range from F8p (Yerkes and Harvard) to G7 (Victoria), while the Harvard classification carries the interesting remark in a footnote, "the spectrum resembles class Ko in the region of *Hβ*". Mt. Wilson assigns type cGo and absolute visual magnitude $-3{\cdot}0$. The spectrum certainly suggests a supergiant with a surface temperature similar to that of the Sun, although the trigonometrical parallax, whose various determinations are not very concordant, suggests that the spectroscopic luminosity may have been overestimated.

In 1946, Popper (**2**) announced that a deep minimum in the light-curve had been accompanied by the development of *TiO* bands in the spectrum and the disappearance of the F supergiant characteristics. It therefore becomes of special interest to publish the results of some observations of this stellar spectrum based on plates taken with the Newall 25-inch refractor at Cambridge in 1939. Further, the publication of Miss Moore's "Revised Multiplet Table" (**3**) permits identifications in conveniently condensed form. The Cambridge plates of 1939 are listed in Table I; designations "T" and "F" refer to the 2-prism and 4-prism plates respectively. The 4-prism plates have dispersions of 34 A./mm. and 13·4 A./mm. at D and *Hγ* respectively, while the 2-prism plates have a dispersion of 48 A./mm. at *Mg* 5183.

TABLE I

Observations of ρ Cass 1939

Date (1939)	h	Plate	Exposure (min.)	Region
Jan. 12	20·6	T 899	85	*Hβ*–D
Jan. 12	22·5	T 900	86	*Hβ*–D
Feb. 5	21·0	F 719	240	4200–4600
Feb. 12	20·2	F 723	180	4200–4600
Apr. 14	20·8	F 764	125	D
Aug. 26	22·0	T 1045	60	*Hβ*–D
Aug. 30	21·5	T 1046	45	*Hβ*–D

The 2-prism plates were taken in the course of a spectrophotometric programme of G5–K5 stars, and ρ Cass was put on the observing list on the basis of the Victoria classification. Interest was aroused by the peculiar spectrum, in particular by the strength of the *Na* D lines combined with a supergiant spectrum similar to α Persei (cF5). The measured equivalent width of $D_1 + D_2$ is 4·9 A., approximately equal to that in a normal K2 star and at least three times that of the width in α Persei. A table of wave-lengths, intensities and identifications of 264 lines measured between 4226 and 6374 A. in this spectrum has been deposited with the Society. The measures of the 4-prism plates were made in duplicate by the writers.

Identifications.—The elements identified are listed in Table II together with the numbers of the representative multiplets as given in C. E. Moore's "Revised Multiplet Table" (1945). If a multiplet contains one or more lines which appear in the spectra as a major contributor it is given in the second column, and in clarendon type if such a line is strong. Multiplets represented only by blended lines contributing 50 per cent or less to the observed intensity are given in the third column.

TABLE II

Elements and Multiplets (R.M.T.) present in ρ Cass 1939
(*Region* 4226–6350 A.)

Element	Main Contributors	Minor Contributors
H	**1**	
*Na*I	**1**, 6	
*Mg*I	2, 11	1, 9
*Mg*II	**4**	
*Si*II	2, 5	
*Ca*I	2, 3, 5, 47	4, 21, 35
*Sc*II	**7, 14, 15,** 23, 24, 26, 28, 29	13, **31**
*Ti*I		?3, 38, 42, 173
*Ti*II	17, 18, **19, 20, 31, 40, 41,** 49, **50,** 59, 60, **69,** 70, **82,** 86, 92, **94,** 103, 104, 113, 114	29, 30, 34, 39, 48, 51, 61, ?68, 93, 94, 115
*Cr*I	1	7, 22, 33, 94, 186, 246
*Cr*II	24, **30,** 31, 43, 44	22, 23, 25, 29, ?39, 50
*Mn*I		16, 21
*Fe*I	2, 15, 16, 27, 37, **41,** 68, **71, 152,** 318, 350, 383, 554, 686, 1180	1, 3, 36, 39, 42, 66, 69, 114, 169, 207, 518, 553, ?649, 687, 691, 693, 828, ?963, ?982, 984, 1065, 1067, 1068, 1087, 1089, 1092, ?1143, 1146, 1163, 1165
*Fe*II	**27, 28,** 32, 35, **37, 38, 42,** 43, **48, 49,** 74	41, 55, 186
*Co*I		?150
*Ni*I	98, ?143, ?162	59, 130, 145
*Y*II	5, 12, 20, 21, **22,** 27, 34	13, 19, 38
*Zr*II	40, 79, 88	95
*Ba*II	1, 2	3
*La*II	8, 24, 25	7, ?36, ?37, 41, 50, 53
*Ce*II	135	3, 81, 108
*Nd*II	?50	49
*Sm*II		14, 37

Wave-lengths.—Following Moore's Catalogue of Radial Velocities, a velocity of −42·6 km./sec. relative to the Sun of ρ Cass was assumed in reducing wave-lengths. This velocity is the mean of concordant determinations at three Observatories, and no report of variation has been recorded, except that since the recent minimum accompanied by the interesting spectral changes, Keenan (**4**), using infra-red material, has reported velocities of −55·8 to −79·0 km./sec. for *Fe* and *Ti* lines respectively.

About 120 relatively unblended lines are available in the Cambridge 1939 material, and when these are sorted according to element and compared with laboratory wave-lengths the residuals from velocity −42·6 km./sec. shown in Table III, with corresponding weights, are derived.

TABLE III

Residual Velocities from unblended lines, ρ Cass 1939
($V_0 = -42{\cdot}6$ *km./sec.*)

Element	Weight	ΔV (km./sec.)
H	1·5	−28
*Na*I	1·5	− 6
*Ca*I	3	+ 1
*Sc*II	6	+ 1
*Ti*II	21·5	− 5
*Cr*II	9	−10
*Fe*I	13	− 3
*Fe*II	18	− 4
*Y*II	7	0
*Zr*II	4·5	0
*Ba*II	1	−17
Weighted mean		− 4·3

The derived mean velocity is therefore −46·9 km./sec. for ρ Cass relative to the Sun in 1939. It would be of interest to determine velocities from plates taken between this date and the 1946 minimum. No certain conclusions can be drawn from the differing velocities derived for various elements in Table III; *H* and *Ba*II, despite their low weights, seem to have larger negative velocities than observational errors permit, the average deviation from the mean, without regard to sign, for *Ti* and *Fe* lines being 6·2 km./sec.

Intensities.—The plates were standardized photometrically and equivalent widths could therefore be made available. However, since the later Cambridge plates were taken with another spectrograph with much lower dispersion, it was not considered that the labour of mass photometric reduction of line intensities would be justified, and visual estimates have been used, with Dunham's visual estimates in α Persei as a standard of comparison.*

It is necessary here to distinguish between the spectral regions where 2-prism and 4-prism spectrograms were available and where consequently somewhat

* Direct comparison with Cambridge plates of α Persei having the same dispersion show that with a few exceptions (notably hydrogen) all lines tend to be stronger in ρ Cass. (As has already been remarked, the D lines are three times as strong in ρ Cass.) Intensities 1 to 3 on Cambridge plates correspond approximately to intensities 5 to 8 on Dunham's α Persei plates, although the former lines are known to be stronger. Thus the positive intensity differences α−ρ seen in Table IV merely reflect differences in scales of visual intensity consequent on different dispersions.

different intensity scales obtain, the 2-prism intensities being systematically smaller by about 1·7 units. The results of the comparison of unblended lines of well-represented elements in ρ Cass is shown in Table IV. The last column gives the weighted average of $\alpha-\rho$ for 2-prism and 4-prism, the former being corrected by 1·7 units. The order represents roughly that of increasing difference between the intensities in the two stars. Small values of $\alpha-\rho$ mean of course that the lines are strong in ρ Cass.

TABLE IV

Unblended lines in ρ Cass
Comparison of intensities with α Persei

Element	4-prism				2-prism				
	ρ	α	$\alpha-\rho$	n	ρ	α	$\alpha-\rho$	n	$\overline{\alpha-\rho}$
*Zr*II	3·2	4·2	1·0	5	...	...	...	0	1·0
*Y*II	4·0	6·3	2·3	3	1·4	4·2	2·8	9	1·4
*Cr*II	3·5	5·2	1·7	6	1·3	4·3	3·0	6	1·5
*Sc*II	6·3:	7·0	0·7:	3	2·7	6·5	3·8	6	1·6
*Ti*II	4·9	7·2	2·3	14	2·3	5·9	3·6	15	2·1
*Fe*II	4·9	7·5	2·6	12	3·3	7·3	4·0	12	2·4
*Fe*I	4·0	6·1	2·1	7	1·3	6·1	4·8	12	2·7

The comparison with α Pers may be carried a stage further in the following way: lines in α Pers with Dunham intensity 5 in the blue or 6 in the green nearly always appear on the Cambridge plates of ρ Cass. By choosing such a criterion for marginal appearance we obtain the list of lines in Table V which appear to be strengthened or weakened in ρ Cass. It will be realized that such marginal cases are less reliable than the comparison of well-defined, unblended lines in the two stars, summarized in Table IV.

The left-hand side of the table lists all those lines measured in ρ Cass which are either unidentified or correspond to lines in α Pers with Dunham intensities less than 4 in the blue or 5 in the green; the right-hand side lists those lines with intensities at least 5 in the blue or 6 in the green which are *not* found on Cambridge plates of ρ Cass.

The material of Table V may be summarized in the statements that:

(1) the 29 lines "strong in ρ Cass" are distributed among elements as follows: rare-earths (7), unidentified (5), *Sc*II, *Y*II, *Ti*II, *Fe*I * (3 each), *Zr*II (2), *Fe*II, *Si*II and *Cr*II (1, each).

(2) the 24 lines "weak in ρ Cass" are distributed among the elements as follows: *Fe*I (14), *Ca*I (4), *Cr*I, *Cr*II (2 each), *Ti*II, *Fe*II (1 each).

We may conclude from Table IV and less certainly from Table V that while the general characters of the spectra of ρ Cass (1939) and of α Pers resemble one another fairly closely, *Fe* lines are definitely weak, while *Zr*II, *Sc*II, *Y*II and rare-earths tend to be strong in ρ Cass. Hydrogen, though providing strong lines, must be included among the elements weak in ρ Cass relative to α Pers.

These conclusions are of interest when compared with the spectrum of ρ Cassiopeiae in 1946 and 1947 dealt with in the subsequent paper.

* Two of the three *Fe*I lines appear on the 4-prism plates in the yellow; one at least is quite doubtfully identified as *Fe*.

TABLE V

Comparison with α Persei : Marginal appearances
(For explanation see text)

Strong in ρ Cass					Weak in ρ Cass		
λ ρ Cass	Intensity ρ	Intensity α	Source	R.M.T.	λ	Intensity α Pers	Source
4256·01	3	2	?*Ce*II, *Sm*II	81, 37	4425·4	5	*Ca*I
4280·33	3n	2	*Sc*II, ...	15, ...	4441·7	5	*Ti*II
4291·58	2	3	*Fe*I	3, 41	4447·8	5	*Fe*I
4322·50	?1	2	*La*II	25	4535·6 }	5W	*Cr*I, *Ti*I
4364·76	1	2	*Ce*II	135	35·9 }		
4403·34	2	3	*Zr*II	79	4552·4	5	*Ti*II, *Fe*I
4463·05	?1	2	*Nd*II	50	4581·4	5	*Ca*I
4484·11	2	3, 2	*Ce*II, *Fe*I	3, 828	4603·0	5	*Fe*I
4512·10	?1n	2W	*Sm*II, *Ti*I	14, 42	4654·6	6	*Fe*I
4539·49	?2	3	?*Cr*II, *Ce*II	39, 108	4691·4	5	*Fe*I
4719·37	1n	2	*Ti*II	59	4771·6	5	*Fe*I, *C*I
4969·12	1	...	?		4903·3	5	*Fe*I
5035·24	1n	3	?*Ni*I	143	4985·4	7	*Fe*I
5055·61	1	3	*Si*II	5	4994·1	5	*Fe*I
5119·44	1	2	*Y*II	20	5139·4	9	*Fe*I, *Fe*I
5132·85	1	3	*Fe*II	35	5142·7	6	*Fe*I, *Ni*I
5192·31	1	4, 4	*Zr*II, *Fe*I	95, 383	5150·9	8	*Cr*I
5211·79	1	3	*Ti*II	103	5237·3	6	*Cr*II
5306·04	1	4	*Cr*II	24	5266·6	6	*Fe*I
5490·72	1n	3	?*Ti*II, *Ti*I	68, 3	5313·7	6	*Cr*II
5544·59	1	1	*Y*II, *Y*II	27	5572·9	6	*Fe*I
5666·8	?1	3	*Sc*II	29	5991·4	6	*Fe*II
5669·40	3	4	*Sc*II	29	6102·7	7W	*Ca*I
5766·3	2	1	?		6122·3	7	*Ca*I
5781·83	1	1	*Y*II	34			
5830·2	1	...	?				
5834·7	?	...	?				
5842·7	?1	...	?				
5862·4	1	3	*Fe*I	1180			
6008·5	1	1	?*Fe*I	?982			

The writers acknowledge with thanks the assistance of members of the Solar Physics Observatory staff in obtaining some of the plates.

Yenching University, Peiping :
1948 *March.*

Solar Physics Observatory :
Cambridge.

References

(1) Prager, R., *Geschichte ü. Literatur des Lichtwechsels der Veränd. Sterne.*
(2) *Harvard Announcement Card*, 752, 1946 June 11; *I.A.U. Circular*, 1050.
(3) Moore, C. E., *Contr. Princeton U. Obs.*, No. 20, 1945.
(4) Keenan, P. C., *Ap. J.*, **106**, 295, 1947.

THE SPECTRUM OF ρ CASSIOPEIAE 1946–1947

A. D. Thackeray

(Received 1948 April 19)

Summary

Ten plates of ρ Cassiopeiae covering the region 4600 to 6600 A. taken during the unusually deep minimum of 1946 have been measured. *Ti* accounts for many absorption lines. *Hα* and *Na* D showed in emission. Some other emission lines are tentatively attributed to *Ti* and are explained in terms of the formation of *TiO* molecules from the *Ti* (^{3}P) state. The weakness of many absorptions is probably due to superposed emission. The absorption lines show a negative velocity of about 40 km./sec. with respect to the 1939 phase.

In 1947 April the *TiO* bands disappeared and the star brightened again nearly to its 1939 state while the supergiant line spectrum which reappeared corresponded to a lower temperature. The absorption velocity resumed its 1939 value. Identifications of multiplets present are listed. Despite the difference in character from the 1939 spectrum, lines of *Y*, *Sc* and *Zr* are again strong, while *H* and *Fe* are weak.

The complex changes do not suggest the presence of two stars. The possible operation of abnormal opacity in the atmosphere due to a deficiency of hydrogen is briefly discussed. The distribution of certain atoms in various stages of ionization is not in accordance with ionization theory and this is tentatively ascribed to the process of dissociation of molecules.

The last stellar spectrum photographed at Cambridge before the war was of ρ Cassiopeiae on 1939 August 30. News of Popper's * announcement of the change to spectral type M was received at the time of the writer's resumption of astronomical work in 1946 August and further Cambridge plates were immediately obtained. A spectrograph with a Wood aluminized grating throwing a high percentage of light into the first order was used, attached to the Newall 25-inch refractor. A 10-inch camera, giving a dispersion of 60 A./mm., enabled the star to be photographed in two hours during its minimum phase. The visual magnitude was estimated on several occasions to be between 6·2 and 6·4, i.e. 1·5 mag. fainter than in 1939. Extraordinarily poor observing weather involved much wastage of time, but nevertheless a useful series of plates was obtained, of which those listed in Table I were selected for measurement.

Plates G148 and 453 were taken with a 20-inch camera and dispersion 30 A./mm., the others with the 10-inch camera.

The star was found to have brightened by nearly a magnitude in 1947 April and the *TiO* bands to have practically disappeared. The brightening, coupled with better weather, enabled some good material to be obtained at this phase. No well-marked changes have been noted since. It is convenient to group the material into two phases: 1946 (August and December) and 1947 (chiefly April). In addition two rather poor plates of 1946 October are available for comparison with August and December. Interest was concentrated primarily on the green-yellow region, including the D lines, while after 1947 April the study was extended

* *H.A.C.*, 752; *I.A.U. Circular*, 1050.

TABLE I

Observations of ρ Cass 1946–1947

Date			Plate	Exposure	Region (A.U.)
	d	h			
1946 Aug.	11	01·0	G2	70	5800–6570
Aug.	17	23·2	5	120	4300–6200
Aug.	25	22·8	6	75:	5250–5900
Aug.	26	21·8	7	70	5450–5900
Aug.	30	02·5	11	125	5164–6500
Oct.	26	01·0	75	135	5200–6100
Oct.	28	19·5	77	40:	6070–6570
Dec.	18	20·2	148	180	5750–6150
Dec.	20	20·4	150	120	4950–6150
1947 Jan.	18	18·7	210	50:	5400–5900
Apr.	9	20·1	284	110	4960–5896
Apr.	12	01·2	288	125	5750–6870
Apr.	15	23·6	293	135	5660–6870
Apr.	19	20·9	294	120	4955–5700
Apr.	22	20·9	297	110	4600–5700
Apr.	27	01·0	300	70	5750–6870
May	30	23·5	320	180	4065–4854
Sept.	21	22·3	416	50	4470–6280
Sept.	26	21·8	418	45	4215–4576
Oct.	23	19·6	453	65	4700–6500

into the blue region. A preliminary brief description of the spectra has been published elsewhere.*

1946 *August.*—The first plates not only confirmed the presence of *TiO* band absorption, but also gave evidence of weak emission at the D lines instead of the strong absorption of 1939. The best plate, G5† of 1946 August 17, displayed the extraordinary characteristic of doubled D2 emission and single D1 (see Fig. 1). Apart from these lines and the *TiO* band-heads it proved to be impossible to identify with certainty any of the numerous features in the complex spectrum. In particular, the absence of strong *Fe* lines (e.g. R.M.T.‡ multiplet 15) which form the most prominent atomic features in the green spectra of M stars was most surprising. Nor was it possible to identify lines with those characteristic of the F-type stage of 1939. It was not until the disappearance of the bands that atomic lines could be confidently identified (1947 April); then it proved possible to trace back some of the stronger features to the earlier phases.

Spectra of M stars taken with moderate or low dispersion have limited value for accurate measurement of wave-lengths and intensities owing to the predominance of complex blends. And in this instance the observed emission at the D lines is a warning that many features may represent a most complicated balance of emission and absorption and that, quite apart from difficulties of measurement, the observed wave-lengths and intensities of lines may lose most of their normal significance. Nevertheless, owing to the extraordinary character of this transitory phase of 1946 it seems worth while to publish a full list of measured

* A. D. Thackeray, *Nature*, **160**, 370, 1947.

† The writer is indebted to Mr W. H. Manning and Mr K. Aly for this plate, obtained on one of the very few nights at this season in which the weather permitted a full exposure. The writer also acknowledges with thanks the assistance of Mr H. E. Butler and other members of the Observatory staff in obtaining some of the other plates.

‡ C. E. Moore, "Revised Multiplet Table", *Contr. Princeton U. Obs.*, No. 20, 1945.

wave-lengths and visually estimated intensities, more especially as some absorptions have no counterpart in the 1939 and 1947 spectra and still defy identification.

FIG. 1.—*Microphotometer tracings of the D lines in ρ Cass* 1946–1947.
Top to bottom : 1946 *August* 17.
1946 *October* 26.
1946 *December* 18.
1947 *October* 23.
The lower two tracings have twice the dispersion of the upper two.

Table II therefore lists average wave-lengths and intensities measured in the August and December plates; the wave-lengths have been reduced to Sun, assuming a stellar velocity of −42·6 km./sec.

TABLE II

Explanation

Wave-lengths reduced to Sun assuming a stellar velocity of −42·6 km./sec.

Wave-lengths given to nearest angstrom unit only refer to features measured only on microphotometer tracings.

Intensity Notation : All intensities refer to absorption unless accompanied by the symbol *e* (emission).

R,V bandhead degraded to red or violet, intensity scale up to 6.
n diffuse absorption.
N very diffuse absorption.
s sharp absorption.
d double.
m absorption maximum in complex structure.

Contributors. Elements, and corresponding multiplet numbers, are given in order of wave-lengths of the various contributors. In a few cases the roman numeral describing the state of ionization of the element is dropped for economy of space; in all such cases the corresponding element is *neutral.*

ρ Cassiopeiae 1946

	August λ	Int.	Oct.	December λ	Int.	Contributors Elt.	Multiplet (R.M.T.)
	4586	6				?*VI*, *TiO*	4,
	87·6	?e 1n					
	4598·7	?e 2n					
	4620·3:	3N				*VI*	4
	28·7	3s				?*TiO*	
	52·3	3(?*V*)					
	4653·6	e 3n					
	4710·3	4				?*Ti*I	75, 203
	11·9	e 5					
	13·2	? R2					
	30·5	? V					
	32·6	?e 1					
	60	R				*TiO*	
	93·6	? V					
	4795·5	?e 1					
*	4911·1:	mN				—, *Ti*I	—, 157
*	24·5:	mN				?*Fe*II, *Ti*I	42, 157
*	33·2	4				*Ba*II	1
	39	m				*Fe*I, *Fe*I	318, 16
	4951·7	R5	×	51·8	R4	*TiO*	
*	5050·9	6				*Fe*I, *Fe*I	114, 16
	59·4	? 2		58·5	3	*Fe*I	1
*				64·2	3	*Ti*I	5
				84·0	m		
	5100	m		00	m		
*				09·1	m	*Fe*I	16, 1
	32·9	? 1					
	37·0	? 1					
*				44·6	9n	?*Ti*I, *Ti*I	109, 4
*	53·5	? 3n		52·1	7n	*Fe*I, *Ti*I	16, 4
	5163·8	R4	×	63·7	R6	*TiO*	
				74·7	2		
*	5203·2	? R		03·5	4	*Cr*I	7
*				09·4	4	*Cr*I, *Ti*I	7, 4
				15·8	2		
				17·8	2		
	18·5	?e 1					
	21·8	? R3					
*	23·6	7	×	24·1	8	*Ti*I, *Ti*I	183, 183
	45·8	3n	×	46·3	4	*Fe*I, *Ti*I	1, 183
*	50·2	3	×	49·7	5	*Fe*I	1
*	54·6:	3	×	54·2	5	*Fe*I, *Ti*I	1, 183
*	63	3		63·4	15n	*Ti*I	183
	67·8	1		68·6	? 2		
*	74·6	4n		74·8	4n	*Cr*I	94
*	82·6	6N		81·7	5	*Ti*I, *Ti*I	74, 156
	85·2	e 5					
				88·9	1	*Ti*I	36
*	5296·5	9N	×	96·1	15	*Ti*, *Cr*, *Ti*, *Cr*	74, 18, 156, 94
	5305·9	R1	×	05·0	? R2	*TiO*	
	10·6	? 2					
				29·4	? V1	? *OI*	12
	5339·9	? 1	×	39·6	? V2		
	42·5	? 1	×	41·7	? V2		
	5349·8	? V2	×	49·7	? V		

	August λ	Int.	Oct.	December λ	Int.	Contributors Elt.	Multiplet (R.M.T.)
	56·6	? 2	×				
				61·9	1	? *Ti*I	35
			?	71·5:	? V1	? *Fe*I	15
*							
	76·1	2	×				
	83·3	2				? *Ti*I	35
	88·7	3				? *Ti*I	35
*	5395·7	1		95·3	4n	*Ti, Ti, Fe*	3, 155, 15
			×	97·9	? V3		
		.		05·6:	4	*Fe*I	15
*				08·4:	8	*Ti*I, *Cr*I	155, 18
*				14·3	2		
				19·3:	1		
	5424·3	3n	×	24·8:	3	—, *Ti*I	—, 3
*	31·8	4n	× ×	33·8	8N	*Mn*I, *Fe*I	1, 15
	36·1:	2				? *O*I	11
	44·1	R5	× ×	44·7	R5	*TiO*	
	51·9	2					
				89·8	3	*Ti*I, *Ti*I	107, 3
			×	90·6	m		
*	5511·4	10N		12		*Ti*I	106
	19·7	1					
*	26·5	5n	×	27·3	3n	*Sc*II, *Mg*I	31, 9
*	33·6	3N	×	34·5	4n	*Ba*I	1
*	44·1	4n	× ×	45·0	2	*Y*II, *Y*II	27, 27
	91·2	? V1	×				
	94·3	? 1					
	5597·7	? R	×			*TiO*	
	5627·6	? 1				? *TiO*	
	49·5	3					
*	56·9	? R4		58·4	? 2	? *Sc*II, *Fe*I	29, 686
*	70·2	2(?V)	× ×	72·1	V5	?*Sc*II, *Si*II, *Sc*I	29, , 12
	81·8	1(?R)				? *Na*I	6
*	5686·8	3n	×	88·9	? V3	*Sc, Na,* ? *Si*II	12, 6 ,
	5705·7	2	×	07·8	? 2		
*	26·0	3		26·7	? 5	*V*I, *V*I	35, 35
	35·0	2				? *Zr*I	4
	42·5	2(?V)					
	44·3	? e					
	46·9	2n	×				
	49·2	? e					
	52·0	3n					
	57·6	R	×			*TiO*	
	71·1	? 1					
	84·1	1					
	5796·4	4	×	97·4	6	*La*II, *Zr*I, O_2	4, 4,
	5809·2	R4	×			*TiO*	
	33·5	? 2n	×				
	37·8	2					
	45·6	R	×			*TiO*	
*	52·2	? V				? *Ba*II	2
	5878·5	? 2		79·4	5	*Zr*I, ? *Ti*I	4, 71
				88·4	5		
	5890·2	e 6?d	×	91·8	en	*Na*I	1
				94·2	4		
	5895·1	e 4	×	96·0	e3	*Na*I	1
	5902·7:	6s				? *Ti*I	71

August λ	Int.	Oct.	December λ	Int.	Contributors Elt.	Multiplet (R.M.T.)
			21·1	? 1	? *Ti*I	72
			40·5	4	*Ti*I	72
48·0	R3		47·0	? R3	*TiO*	
			65·0	5	*Ti*I	154
75·0:	3n	?	77·5	7N	?*Fe*I, *Ti*I	959, 154
			84·8	? 1		
		?	91·2	2	*Ti*I	154
5996·3	R3		97·1	? R1		
			40·0	? 1		
6061:	1n(?V)					
		×	64·2	12n	*Ti*I, ?*ScO*	69,
			66·4	? R2		
			71·6	5n	?*ScO*	
76·9	? R	×				
			79·9	9n	?*ScO*, *V*I	, 34
			83·4	10n	*Ti*I	69
6089·0	3n	?	90·2	7n	*V*I, *Ti*I	34, 238
		×	02·6	11n	*Ca*I	3
			11·4	5	*V*I	34
			15·3:	? 2		
6119·6	3n		20·7	8	*V*, *Ti*, *Ca*	34, 153, 3
25·3	4n		26·2	7	*Ti*I, *Zr*I	69, 2
35:	3n(?V)		35·8:	10n	? *YO*, *V*I	, 34
			41·9	4	*Ba*II	2
45	R4	×	47·9	R4	? *YO*, *TiO*	
52·6	5					
			54·5	3	? *Si*I	29
58·1	6		57·4	R4	?*O*I, *TiO*	10,
63·5	3				—, ?*Ca*I	—, 3
72·1	? 1				?*TiO*	
6192·7	R2s					
6213:	3n				?*TiO*, *V*I	, 20
20·0	2				?*TiO*	
24·6	? 3				?*V*I	20
36·8	6n				?*Si*I	28, 27
44·7	? 2				?*Si*I, *V*I	28, 27, 19
48·3	? 2					
58·2	2s				*Ti*I, *V*I	104, 19
6283·0	2					
6309·9	2					
44·8	3n					
6348·9	? 3				?*TiO*	
6410	? R1				?*TiO*	
31·1	2				*Fe*I	62
6437·0	3				*Ca*I	18
6561·3	e 5				*H*	1

In addition, features measured on the October plates, at closely corresponding wave-lengths, are marked × in the column headed "Oct.". An asterisk in the first column denotes correspondence with a strong feature in 1947 April. Under "Contributors" are given the elements and R.M.T. multiplet numbers to which the line or band is mainly ascribed; many of these identifications must be regarded as most tentative, being suggested, as mentioned above, largely in the light of subsequent behaviour.

Elements present in 1946.—It is convenient to discuss the various elements in order of increasing atomic number.

H.—The region of $H\alpha$ is not covered adequately, but a fairly strong, out-of-focus, bright line was measured on G2 (August 11) at 6561·3, and a weaker one at 6562·5 on G77 (October 28), both of which are presumably to be ascribed to $H\alpha$. No sign of $H\beta$ emission or absorption was found; but plates are often rather weakly exposed here.

He.—No evidence, though the region of D3 (E.P.20·9) is very well studied.

Li, Be, B, C, N.—No evidence.

O.—*O* is clearly present in molecular *TiO*. *ScO* is also suspected at 6064, 6072, 6080.

Atomic *O* probably contributes weakly with multiplets 10, 11, 12 (E.P.10·7). Multiplet 1 (E.P.9·1) is known to have been strong during the minimum of ρ Cass.* The appearance of these high excitation lines together with oxide absorption probably indicates abundant oxygen.

F, Ne.—No evidence.

Na.—D emission has already been noted. The structure of the lines is most complicated and seems to vary from plate to plate. Absorption of varying strength is found chiefly on the shortward border increasing gradually with time. Multiplet 6 seems to be present weakly. (See Fig. 1, and Fig. 2, Plate 6).

Mg.—Multiplet 9 blended with *Sc*II is a strong and persistent feature. The "b" triplet (2) is masked by *TiO*, but reappears weakly in 1947 April.

Al.—No evidence.

*Si*I.—28, 29 (E.P.5·6) are probably present. 9 can hardly contribute to λ5797 owing to the absence of other members of this multiplet and of 10, 11. Keenan found some lines of multiplet 23 abnormally strong.

*Si*II.—The pair (2) (E.P.8·1) appears to be very strong on a weakly exposed October plate. The lines must have been much weaker in August. They are weak and diffuse in 1947 April.

The strength of *Si* in ρ Cass was noted by Miss Burwell † before this minimum and the presence of these high excitation lines again probably indicates high abundance.

P, S, Cl, A, K.—No evidence.

*Ca*I.—Multiplet 3, partially identified in August, strengthened considerably in December and remained thereafter. 20 is also represented. The non-detection of strong lines due to 21, 22 until 1947 is noteworthy.

*Sc*I.—12 may contribute to two blended lines (very strong in 1947).

*Sc*II.—31, blended with *Mg*, is strong. 29 is also present but in nothing like the extraordinary strength shown in 1947 April.

The oxide appears to be present. (See under *O*.)

*Ti*I.—This element provides most of the atomic features as well as the molecular oxide absorption. Identification in most cases seems certain, especially when compared with the 1947 spectrum, though in many cases the intensities are capricious, presumably due to emission partially filling in the absorption.

Multiplets contributing to measured lines include the following numbers: 3, 4, 5, 35, *36*, 69, 71, *72*, 74, 75, 106, *107*, *109*, *153*, 154, 155, 156, 157, 183, 203, 238. All but those in italics were detected in August. The very strong multiplet 38 was masked by *TiO*.

* Keenan, *Ap. J.*, **106**, 295, 1947.
† Burwell, *Ap. J.*, **88**, 281, 1938.

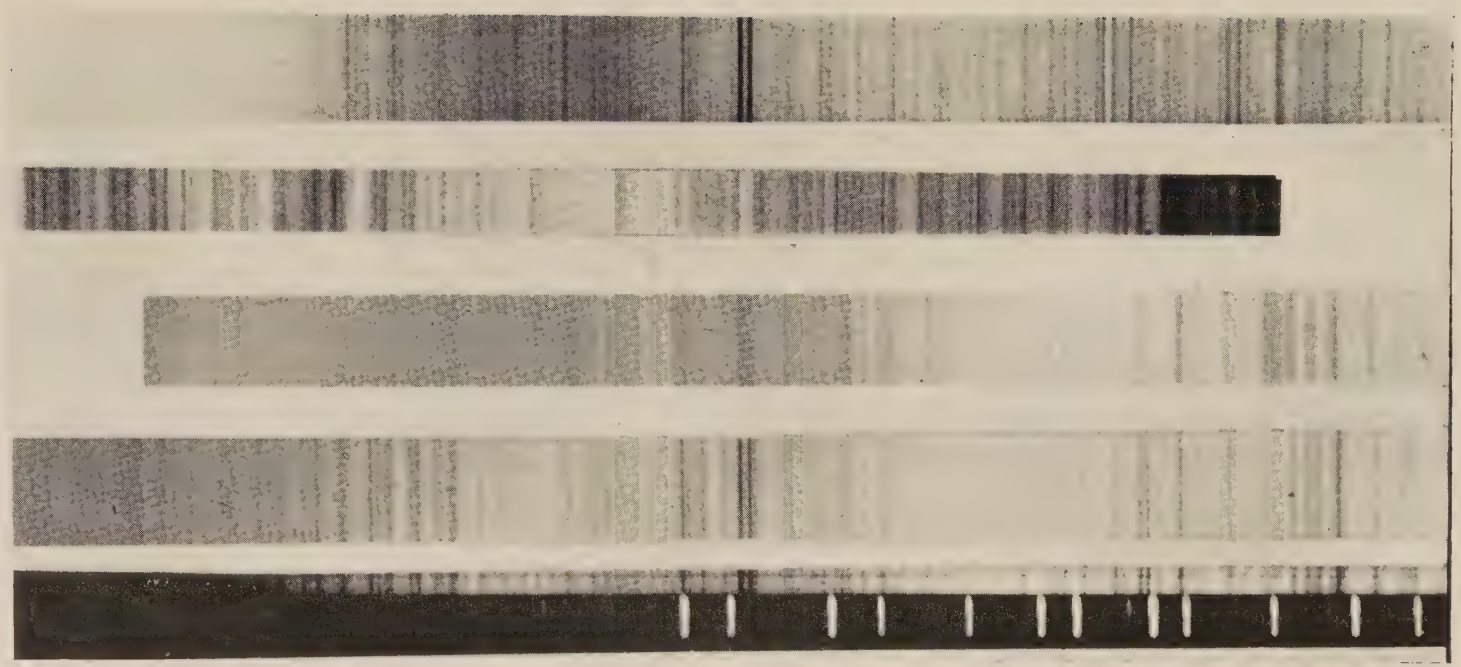


PLATE I.

FIG. 2.—*D lines in ρ Cassiopeiae.*

Top to bottom : *ρ Cass* 1939 (*4-prism*).
ρ Cass 1946 *August* 17.
μ Ceph (*cM2*).
ρ Cass 1947 *October* 23.
ρ Cass 1947 *October* 23 (*unwidened*).
Ne comparison.

1946 *and* 1947 *spectra have the same dispersion* (*grating*).

A. D. Thackeray, *The Spectrum of ρ Cassiopeiae, 1946–1947.*

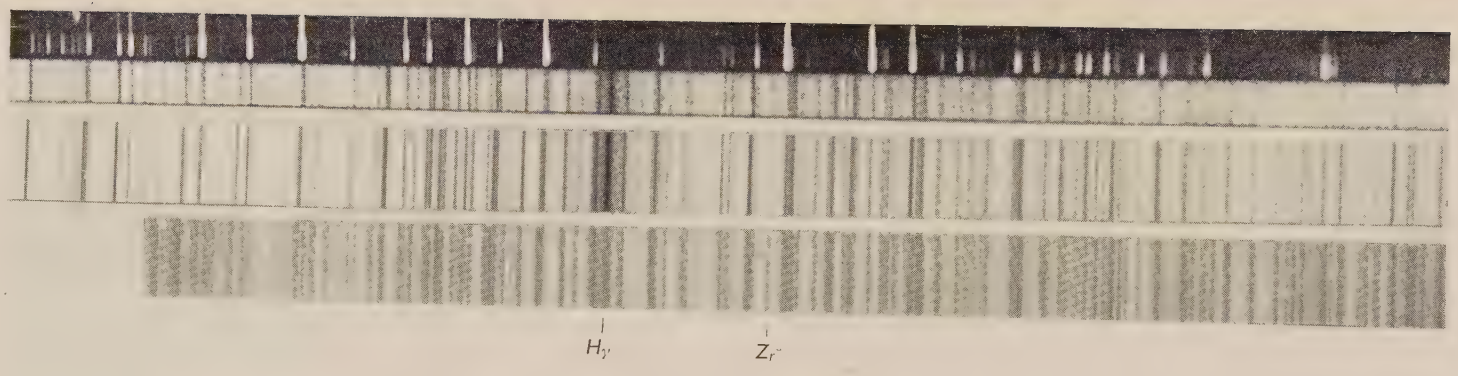


PLATE II (*a*).
Top to bottom : Fe comparison.
α Persei (unwidened).
α Persei (widened).
ρ Cass 1939 *(widened).*

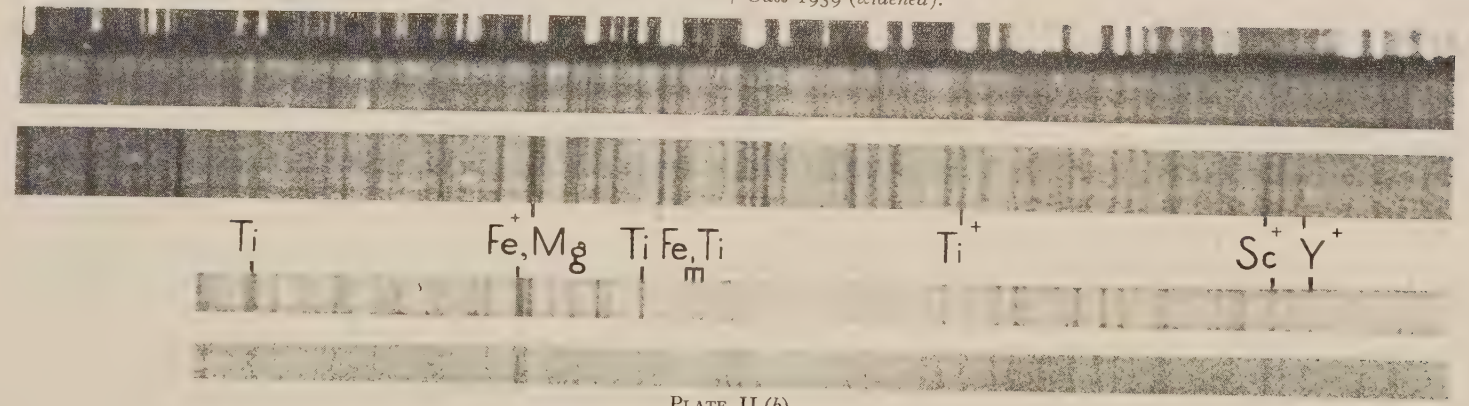


PLATE II (*b*).
Top to bottom : Fe comparison (4950-5650).
β Caml (*cG*0).
ρ Cass 1939 *January.*
ρ Cass 1948 *April.*
δ C Maj (*cF*8).
The upper three spectra have prismatic dispersion.
The lower two have normal (grating) dispersion.

A. D. Thackeray, The Spectrum of ρ Cassiopeiae, 1946–1947.

Particularly strong absorption, attributable mainly to *Ti*, was found at λλ 5224 and 5263 (183), 5296 (74, 156), 5511 (106), 6064 and 6083 (69).

For emission lines possibly due to *Ti* see later.

*V*I.—34, 35 are believed to be present; despite the prominence of 19, 20 in April, these multiplets could not be detected in 1946.

*Cr*I.—7, 18, 74, 94 present in medium strength.

*Mn*I.—A strong line at 5432 may be due on its shortward side to resonance absorption by this element. Otherwise no evidence.

*Fe*I.—The weakness of the lines due to this element is extraordinary. The leading diagonal lines of multiplet 15 are not detected, *with the possible exception of the two weakest*, 5405, 5433, which provide moderately strong lines in December. The satellite lines are lost in complicated structure but become well marked in 1947 April. The satellite lines of multiplet (1), blended with *Ti*, are seen as clear lines despite the absence of other members. If we are correct in ascribing the measured lines λλ 5405, 5433 to (15) and λλ 5246, 5250, 5254 to (1) as an important contributor, we have here remarkably interesting cases of the strengthening of weak lines in multiplets which were measured spectrophotometrically by the writer * in α Orionis and other M stars and which is also evident in many weak lines in the sunspot spectrum.† The appearance of the phenomenon to a marked degree in the 1946 stage of ρ Cass is almost certainly to be explained in terms of emission filling in the absorption contours leaving the strongest lines most weakened. Whether the emission and absorption are to be visualized as taking place in different parts of the atmosphere, or whether it is a true "interlocking" effect taking place on an atomic scale cannot be decided on the present evidence. But the behaviour of the *Fe* lines in ρ Cass certainly emphasizes once again how far the formation of absorption lines in the integrated light of an M-type atmosphere may deviate from the conditions necessary for a classical "curve of growth".

Ni, *Co*, *Cu*, *Zn*, . . . *Rb*.—No evidence. Keenan found the resonance lines of *Rb* in the infra-red.

Sr.—An absorption at 4607 in August may correspond to the resonance line (2).

*Y*II.—Moderately strong absorption at 5545 seen August, October and December apparently corresponds to multiplet 27 which is well identified in the 1947 plates. No evidence of *Y*I, though strong lines are accessible.

*Zr*I.—There is evidence of (2) and (4). The latter is unduly strong if it provides the main contribution to λ 5797. No evidence of *ZrO*.

*Ba*I.—It is of interest to find λ 5535, apparently the resonance line of *Ba*I, well marked at all times in 1946 and 1947; in 1939 a line was measured at 5534·46 but this is attributed to *Fe*II (55); this latter identification will not serve for the minimum phases.

*Ba*II.—(1) is well marked, (2) is doubtfully present but strengthened enormously in 1947.

Rare-earths.—Material inadequate to judge presence or absence of these elements.

Velocities (1946).—Many of the 1946 identifications are so doubtful that it is scarcely practicable to attempt a velocity measurement at this phase. The

* *M.N.*, **94**, 99, 1933.

† C. E. Moore, *Ap. J.*, **75**, 298, 1932. See Dorothy N. Davis (*Ap. J.*, **106**, 34, 1947), who finds the same phenomenon of strong satellite lines in β Pegasi.

best that can be done is to compare wave-length measures of features common to the 1946 and 1947 plates. In this way we find for the difference in velocity δV (1946–1947).

	August		December	
Element	Weight	δV	Weight	δV
Ti	6	−44	16	−42
Fe	6	−39	7	−36
Others	13	−39	14	−28
Mean	25	−40	37	−36

As the 1947 velocity was measured as −45·3 km./sec. (see below), the 1946 corresponding velocities appear to be −85 and −81 km./sec. This is in reasonable accord with Keenan's results for the infra-red.

In view of the known presence of emission it is doubtful whether the large negative velocity of 1946 is to be interpreted as absorption by approaching gases. It might equally well be attributed to the effect of receding gases in emission filling up the longward contours of the absorption lines. Merrill *, however, has brought forward good reasons for supposing that in late type variables absorption lines usually give a good guide to the stellar velocities.

Intensity Variations in 1946.

(a) *Bands.*—The *TiO* bands were observed to be weaker than in β Pegasi (M2) on all Cambridge plates from August onwards. The discontinuities $\Delta \log I$ at various band-heads were measured spectrophotometrically and suggest a slight increase up to December, as shown below:

TiO band-heads in ρ *Cass*, $\Delta \log_{10} I$

Date	4955	5163	5446	5597	5759
1946 Aug. 17	·18	·23	·20	·09:	·08
1946 Oct. 26		·25	·23	·10:	
1946 Dec. 18					·12
1947 Jan. 18					·11
1947 Apr. 22		...	< ·12	...	...

A number of discontinuities were measured as possible band-heads; one at λ5350, degraded to the violet, was measured on most 1946 plates, and weakened as time went on without being replaced by any strong absorption line. Another, measured as a head at 5657, degraded to the red, also weakened through 1946; but this feature may have been caused by *Sc*II line absorption which became strong in 1947.

(b) *Emission.*—The D lines and *H*α have been identified.

The D lines were measured at their strongest in August to be 20 per cent above the neighbouring spectrum, 7 per cent in October, and still weaker, with strong shortward absorption, in December. D2 was measured as double, with D1 single, on plate G5 only, but the former line always appeared somewhat wider.

The two other strongest emission lines on G5 were at 5285·2 and 4711·9 (out of focus), both with strong shortward absorption. The absorption borders of these lines are ascribed to *Ti*, multiplets 74, 156 for λ5285 and 75, 203 for λ4711. It is perhaps significant that 74, 75 both involve the lower state a^3P which is

* Merrill, *Spectra of Long Period Variable Stars* (Astrophysical Monograph), p. 65, 1940.

believed to be a dissociation product of the *TiO* molecule. Consequently the neighbourhood of the other emission lines was searched for possible identification with other *Ti* lines. The table below summarizes the position.

λ (G5)	Em. Int.	Nearest *Ti* (lab)	Lower state
4587·6	? 1n	86·95	z^3F^0
4598·7	? 2n	98·99	c^3P
4653·6	? 3n	55·71	c^3P^0
4711·9	5	10·19	a^3P, a^3D
4732·6	? 1	33·43	a^3D^0
4795·5	? 1	96·21	c^3P
5218·5	? 1	19·70	a^3F^0
5285·2	5	84·38	a^3P^0

It is perhaps easier to account for emission lines having a common *upper* state. But some cases of *lower* states being favoured have also been reported.* The suggestion that the states a^3P and c^3P of the *Ti* atom were favoured for emission lines in ρ Cass therefore deserves consideration, although the identifications in the above table can only be offered with great reserve. It may be mentioned here that just as emission lines originating from a common *upper state* demand a selective mechanism of *excitation* of atoms into that state, so emission lines which terminate on a common *lower* state demand a selective mechanism of removal of atoms from that lower state. The formation of *TiO* molecules from free *Ti* and *O* atoms may provide such a mechanism. The dissociation product of *TiO* being *Ti* (^3P) + *O* (^3P), the *formation* of *TiO* is liable to lead to the removal of *Ti* atoms from the ^3P state, and provided this process occurs sufficiently rapidly to outweigh the supply of atoms in the ^3P state by collisions, emission lines having ^3P as their lower state will appear.

The only other features appearing in Table II as possibly emission occurred at 5744 and 5749 on plate G11 only, and these were doubtfully recorded; more probably they represent spaces between diffuse absorption of unknown origin.

(c) *Line absorption.*—Generally speaking, line absorption emerged and strengthened gradually from August to December, and most of the stronger lines have their counterparts in 1947. The lines were ill-defined and diffuse in 1946 August, became clearer in December, and in April resumed their sharp supergiant character of 1939. However, while α Persei (cF5) forms a reasonable comparison for the 1939 phase, in 1947 μ Cephei (cM2) compares well in the red, apart from the absence or weakness of *TiO*,—in the green β Caml (cG0) is not very dissimilar except that ρ Cass has much stronger lines. It is exceedingly difficult to find any comparison star which matches at all wave-lengths with ρ Cass (1947). The temperature corresponding to the line spectrum seems to vary from 5500 deg. in the green to 3500 deg. in the red.

The identification of λ5797 which was prominent throughout 1946 remains most unsatisfactory. Its equivalent width was measured as 0·5 A. in August and December. Moreover, in contrast to the strengthening of nearly all lines in 1947 April, the equivalent width was found to remain at 0·5 A. in 1947 April and October.

The most reasonable identification for this line seems to be *Zr*I (4). The other lines 5879·8 and 5735·7 were apparently present in August. But 5879 should be stronger (its equivalent width in 1947 October was measured as 0·54 A.),

* Merrill, *Ap. J.*, **106**, 280, 1947.

while 5735 was not measured as a line after August. It is worth mentioning that one of Merrill's unidentified interstellar lines lies near here at 5797·13, but in no case has such a large equivalent width been measured; nor do the Cambridge plates of ρ Cass show any lines which may be identified with others of these diffuse interstellar lines.

1947.—374 absorption lines have been measured on the 1947 plates between 4065 and 6830 A. The best material refers to the region longward of $H\beta$; the $H\gamma$ region is covered by two plates only, G320 and 418.

Since the spectrum had resumed a fairly simple character, practically freed from molecular bands, in which atomic lines could be identified fairly readily, it has been decided to present the results in the condensed form adopted for the preceding paper which describes the 1939 spectrum; i.e. the elements and multiplets identified as present are listed in Table III, while the detailed results have been deposited with the Society.

TABLE III

Multiplets present in ρ Cass 1947. (4065–6830 A.)

Element	Major Contributors	Minor Contributors
*H*I	**1**	
*Na*I	**1**	6
*Mg*I	2	1, 8, **9,** 14
*Mg*II		4
*Si*I	?29	?9, ?10
*Si*II	2	
*Ca*I	**1, 2, 3,** 18, **21,** 32, 47	4, 5, 19, 22, 33, 48, 49
*Sc*I	2, **12,** 13, 19	3, 4, 14, 16
*Sc*II	7, 19, 23, 28, **29,** 30	14, 15, 26, **31**
*Ti*I	1, 2, 3, **5, 6,** 7, 8, 35, 36, 37, **38,** 48, 49, **53,** 69, 72, **102, 104, 106,** 110, 144, 145, 148, 154, **157,** 158, **183,** ?229, 240, **260,** 279	4, 9, 42, 43, 44, 74, 102, 103, 107, 109, 112, 113, 146, 153, 155, 156, 197, 201, ?229, 238, 249, 261, 265, ?269, ?293, ?309
*Ti*II		19, 20, 31, 41
*V*I	**3,** 4, **5,** 19, ?20, 22, 31, 34, 35	6, 27, 37, 86, 107
*Cr*I	?1, 18, 22, 60, 94, 201	?2, 7, ?10, 21, 32, 186
*Cr*II	**30**	26
*Mn*I	4	1
*Fe*I	1, 2, ?14, **15,** 16, 36, 41, **43,** 66, 109, 111, 168, 169, 207, **268, 318,** 553, 686, 687, 816, 1165, 1258	3, 13, 37, 38, 42, 62, 68, 114, 518, 554, 959, 1066, 1067, 1195
*Fe*II	**40, 42,** 46, 74	27, 48, 49
*Co*I		?37
*Ni*I	43, ?45, 59, 69, 68	46, 48, 57
*Cu*I	2	8
*Y*II	**5, 14,** 20, **22,** 26, 27, 38	**1,** 13, 21
*Zr*I	4, ?25, 43, 45	2, 5, ?26, 44
*Zr*II		15, ?41
*Sr*I		2
*Sr*II		1
*Cs*I		?2
*Ba*I	**1,** 6	
*Ba*II	**1, 2**	
*La*II	2, 37, **75,** 91	4, 7, 8, 19, 82
*Eu*I		?1
*Eu*II		?1, 4

Notes on Table III follow, element by element.

*H*I is represented by medium strong *Hα*. *Hβ* is *not detected*, while *Hγ* may contribute to a blended line at 4341·8. The weakness of *Hβ* and *Hγ* is of interest, especially when the relative weakness of the *H* lines in 1939 is recalled.

Na.—The D lines are strong in absorption—spectrophotometric measures in April show them to be stronger than in 1939, but in 1947 October, the equivalent width $D_1 + D_2$ is 4·7 A., practically the same as in 1939. The subordinate pair (6) is also present.

*Mg*I.—The "b" triplet is blended and apparently rather weak. The blend of (9) with *Sc*II (31) is strong, as in 1946.

*Mg*II.—A very strong line at 4480·8 is attributed to *Ti*I (146), *Cu*I (8), *Mg*II (4).

*Si*I and II.—A few lines are present, weaker than in 1946.

Ca.—The intercombination resonance line (1) is strong—rather stronger than *Hα*. Many more lines due to this atom are present than in 1939.

*Sc*I and II.—Lines due to both *Sc*I and II are well represented. *Sc*II (28), (29) and *Sc*I (12) dominate the yellow region with remarkably diffuse lines.

*Ti*I and II.—Extremely numerous lines due to the neutral atom, especially from states of low E.P. *Ti*II is represented by only a few lines. An emission border is seen to the red of 4501·4, attributed to a blend of *V*I and *Ti*II.

*V*I.—Well represented, though conditions were unsuitable for the appearance of any lines due to this element in 1939.

Cr.—The resonance lines (1) are obliterated in heavy structure, but the atom contributes to many other observed lines. *Cr*II also contributed some strong lines.

*Fe*I.—Lines are not quite so numerous as *Ti*I; *Fe*II is better represented than *Ti*II, though much weakened compared with 1939. In 1947 October the *Fe* lines appear somewhat stronger.

*Co*I.—One doubtful contributor to a blend.

*Ni*I.—Presence established by several lines.

*Cu*I.—Presence rather doubtfully established.

*Y*II.—The strength of the lines due to this ion is one of the most remarkable features of the 1947 spectra of ρ Cass. (27) is completely present, thus lending support to the identification of the line at 5545 seen in 1946 with two close members of this multiplet. The only two accessible members of (26) are also easily detected. These two multiplets, with lower state a^3P seem to be inordinately strong.

No lines due to *Y*I have been detected, though many low-level lines are accessible.

*Zr*I and II.—Both represented, the ion only weakly; in 1939 only the ion was detected.

*Sr*I and II.—Resonance lines due to neutral atom and ion contribute to blends.

*Cs*I.—The two resonance lines (2) may contribute to strong blends; thus 4554·1 (7 ?d) is attributed to *Ba*II (1), ?*Cs*I (2), *Ti*I (42); 4593·1 (9) is attributed to ?*Cs*I (2), ?*Eu*I (1), *V*I (4). Intensity and wave-length demand a blend shortward of the *V*I line.

*Ba*I and II.—Both represented by strong resonance lines. 6495·9 is by far the strongest line measured to the red of *Hβ* and is probably mainly due to *Ba*II (2), though *Ca* (18) and *Fe* (168) contribute. In this respect ρ Cass resembles some other M stars, e. g. VV Cephei and μ Cephei.

*La*II.—Several lines present.

*Eu*I and II.—A very strong line at 4205·2 is attributed to *Y*II (1), *Eu*II (1). The neutral atom may contribute to a blend (see under *Cs*).

Velocities (1947).—124 lines were selected as relatively unblended (over half being ascribed to *Ti* and *Fe*). When these lines were grouped according to element and average velocities formed, the following values of δV (deviation from −42·6 km./sec.) were found:

Element	Weight	δV
*Ti*I	25·5	+1·2
*Fe*I	27·5	−1·3
Others	45	−5·6
All	98	−2·7

Velocity relative to Sun −45·3 km./sec.

It thus appears that by 1947 April, after disappearance of the bands, the apparent velocity of ρ Cass had returned to the 1939 value of −42·6 km./sec. within the error of measurement.

The material is scarcely adequate to discuss apparent differences in velocities of various elements; but it is worth recording that the largest deviations from the average δV are shown by *H* −33 (1), *Na* −27·5 (2), *Ba*II −13·7 (3)—the relative weights are given in brackets. Note that the two largest deviations occur for the two elements which were most certainly identified in emission in 1946.

Intensities (1947).—Even after the disappearance of the bands the lines are too numerous for accurate measurement of intensity. An attempt has been made to use the visual estimates, as in the comparison with α Persei in the preceding paper, the basis of comparison with the 1947 spectra being D. N. Davis' * visual intensities in α Scorpii (cM2). As the intensity scale in α Scor appears to be far more open than that of α Pers, in this instance we use directly the ratio of intensities ρ Cass/α Scor. The average ratio for various elements, together with corresponding weights, are shown in the accompanying table:

*Y*II	*Ba*I	*Sc*II	*Zr*I	*Ti*I	*V*I	*Ca*I	*Ba*II	*Cr*I	*Na*I	*Fe*I	*Ni*I	*H*
0·78	0·55	0·39	0·38	0·36	0·35	0·29	0·28	0·26	0·23	0·19	0·15	0·07
9	2	9	6	50	11	18	6	7	4	49	5	2

The elements are arranged in order of decreasing ratio.

If temperature and pressure in the atmosphere of ρ Cass (1947) and α Scor are taken to be the same, the above order may indicate roughly the order of decreasing relative abundance of the various elements in ρ Cass. It is interesting to find *Y*II, *Sc*II and *Zr*I again among the "abundant" elements, as in the completely independent comparison with α Persei in 1939; further, the 1939 comparison was based on *Zr*II lines, while the 1947 comparison uses *Zr*I lines.

Among the "deficient" elements we find again *H* and *Fe*, to which *Ni* is now added, lines due to this latter element failing to appear in sufficient quantity in the 1939 condition.

Physical Discussion.—It is perhaps premature to attempt to piece together the strands of observational evidence and interpret them in a unified picture.

* D. N. Davis, *Ap. J.*, **89**, 41, 1939.

But it is interesting to follow up one or two general arguments. Let us take the surface temperature and absolute magnitude to be

in 1939	$T_1=6000°$,	$M_1=-3{\cdot}0$
while for 1946	$T_2=3100°$,	$M_2=-1{\cdot}5$
and after 1947 April	$T_3=4000°$,	$M_3=-2{\cdot}5$

Then applying the black-body formula* for the radii (in units of the Sun's radius) at the three phases

(1939)	$R_1=36$
(1946)	$R_2=152$
(1947)	$R_3=91$

It is doubtful whether the star is in a sufficiently steady state to apply the black-body formulae, especially to the transitory 1946 phase, but the star seems to have settled down in phase 3 with relatively little change for one complete year. If this phase persists, it appears that here we have a star which has managed to expand its atmosphere more than twofold in the course of a year or two. Yet such a change has not been accompanied by any nova-like catastrophe. It is true that the absorption velocities in 1946 suggest a velocity of approach of 40 km./sec. compared with 1939 and 1947; further, the D lines and λλ 5285, 4711 showed contours analogous to P Cygni emission, suggestive of a mild nova. But this was during light-minimum. Unfortunately, the Cambridge series started too late during the minimum to be sure that the negative velocities were characteristic throughout this phase.

It is clear that the minimum and the accompanying spectral changes do not arise from an eclipse like that of the Zeta Aurigae system. There is nothing in these observations to suggest that more than one star is involved. The light-variations are more of the irregular type shown by α Orionis or long-period variables, but there might well be obscuration by circumstellar clouds, possibly of interstellar origin, of the type implicit in the "veil theory". If it can be imagined that such clouds could obscure a large portion of the disk of a supergiant as seen from the Earth, no radical internal change of stellar structure is required to account for the 1946 phase. The D line emission could be attributed to an outer chromospheric envelope less involved in the eclipse; the central intensities of the D lines as measured in 1939, and corrected for resolving power, amounted to about 25 per cent of the continuous spectrum, and a diminution of 1·5 mag. which left the core of the D lines unaffected would therefore cause complete extinction of the lines; the weakening of other absorption lines could be attributed to a similar circumstance.

If, however, the 1947 phase represents a more or less permanent readjustment to a larger radius it is hardly conceivable that an external eclipse has taken place and we should look to some deep-seated physical change, possibly due to exhaustion of some nuclear source of energy. This is perhaps the most likely explanation. The composition of the atmosphere appears to be definitely abnormal, as judged by the weakness of hydrogen and iron lines and the strength of *Sc*, *Y*, *Ti* and *Zr*.

However, the conditions of excitation must also be regarded as peculiar and interpretation of intensities in terms of abundance must be offered with caution. The presence of emission filling in absorption has already been stressed. Further, the intensities of lines of various atoms in different stages of ionization

* Russell, Dugan and Stewart, *Astronomy and Astrophysics*, Vol. 2, p. 738.

do not conform to expectations based on ionization theory. If we tabulate the elements in order of increasing first ionization potential we find the following state of affairs as regards strength of lines due to neutral and once ionized elements; +, to + + + indicates presence in increasing strength, O indicates strong accessible lines not detected, ind. means no evidence:

Neutral and once ionized Atoms in ρ Cass, 1947 April

Element	First I.P.	Second I.P.	Neutral	Ion
Na	5·1	47·1	++	ind.
Ba	5·2	10·0	+	++
Ca	6·1	11·8	++	ind.
Y	6·5	12·3	O	++
V	6·7	14·1	++	?
Cr	6·7	16·6	+	+
Sc	6·7	12·8	+	+++
Ti	6·8	13·6	+++	+
Zr	6·9	14·0	+	+
Mn	7·4	15·6	?+	ind.
Mg	7·6	15·0	+	?+
Fe	7·8	16·2	+	++
Si	8·1	16·3	?+	+
H	13·5	...	+	...

Attention is drawn to the fact that *Y* and *Sc* are unduly strong in the once ionized state, while *Ti*, with comparable ionization potential, predominates in the neutral state. No explanation in terms of an ultra-violet anomaly at the various ionization limits is forthcoming. An alternative suggestion offers itself in terms of dissociation of molecules; but at the present stage it can only be regarded as a working hypothesis; the abundance of *Ti* is apparently associated with the presence of *TiO* molecules dissociating into $Ti + O$, and the abundance of *Sc* and *Y* may equally well be associated with the presence of *ScO* and *YO* dissociating into $Sc^+ + O^-$ and $Y^+ + O^-$ in conditions of low pressure. Laboratory experiments on the dissociation of these oxides would be interesting.

One further complication regarding the interpretation of line-intensities in this spectrum remains to be mentioned—the *continuous absorption* at various wave-lengths. Since the 1946 minimum Cambridge material has shown a spectrum which not only combines high-excitation and low-excitation lines, as Keenan has stressed, but which also varies in type from late K in the red to late G in the green. Other observers have classified the star into different types at different wave-lengths without detecting one definite boundary between the types as in the normal "composite spectrum". Is it possible that here we are dealing with an extended atmosphere whose opacity varies abnormally with wave-lengths and with time in such a way that the temperature of the "photosphere" varies from late F to K or M not only with time but also, at a given epoch, with wave-length? Such an opacity could account for the difficulty of matching all regions of the spectrum with one star. The colour of the star has been compared with that of 56 Pegs (G8) and remained practically unchanged through the minimum.*

* Popper, *A.J.*, **52**, 129, 1947.

The writer is inclined to invoke a *deficiency of hydrogen* as the fundamental property underlying the peculiarities of ρ Cassiopeiae.*

In this way we might account for

(1) the weakness of the hydrogen lines;

(2) the existence of an extensive and abnormally transparent atmosphere in which high excitation lines of *O* and *Si* are discernible;

(3) the dominance of another source of opacity replacing that of *H* and H^- in the normal stellar atmosphere; during minimum the source of opacity was probably mainly molecular, but O^- must be regarded as a likely alternative to H^- after the minimum;

(4) the formation of *TiO* molecules at what appears to be a rather higher temperature than usual; Russell's † curves suggest that when hydrogen is deficient the free oxygen, available from the suppression of *OH* and H_2O, will largely be taken up to form *TiO*;

(5) possibly, the initiation of an unstable phase in the star's evolution due to local deficiencies in the star's interior of hydrogen as the main source of energy.

Solar Physics Observatory,
Cambridge:
1948 *April.*

* The hydride absorption reported by Greenstein and Swings in *I.A.U. Circular*, No. 1117, as present in ρ Cass on a few nights in 1947 November must be regarded as a transitory phenomenon not necessarily inconsistent with a relative deficiency of hydrogen. No corresponding absorption is found on the Cambridge plate of 1947 October 23, nor in 1948 April.

† *Ap. J.*, **79**, 329, 1934.

THE CORONAL EMISSION SPECTRUM

R. v. d. R. Woolley and C. W. Allen

(Received 1947 December 19)

Summary

The ionization in the corona is computed by supposing it to be determined by a balance between ionization by electron impacts and photoelectric recombinations. The results for an electron temperature of 10^6 degrees agree with experiment. An estimate of the abundance of elements in the corona, from coronal line intensities in relation to the electron scatter continuum, gives a hydrogen to metal ratio of 6000: 1 by numbers. The collision excitation cross-section of *Fe* XIV 5303 A. is found to be $0{\cdot}26\pi a_0^2$ by a method which does not assume the abundance of *Fe* XIV. The coronal emission in discrete lines, in bound-free continua, and in free-free continua, is separately computed. It is found that the principal lines of *Mg* X at 610 A. are strong in emission. It is suggested that the emission from the chromosphere in the Lyman continuum is stronger than the coronal emission, and the effects on the ionosphere of these spectra are briefly reviewed.

The identification of the coronal emission lines by Edlén* has secured general agreement for the coordination of a large number of observational facts into one important conclusion, namely that the corona is at the high temperature of approximately one million degrees.† This can be regarded as firmly established, and the conditions derived from it are now being applied to problems that are further afield, such as solar radio-noise ‡, and the excitation of the ionosphere.§ We therefore inquire whether further study of the mechanism of coronal emission can lead to reasonably accurate information of the type required by these and similar problems.

1. *Ionization of the Corona.*—We shall start by calculating the degree of ionization of the corona. The temperature, judged by the agitation of the molecules, is taken to be of the order of 10^6 degrees, while the radiation density corresponds to a much smaller temperature. Saha's relation, which is valid for thermo-dynamical equilibrium, cannot be expected to apply, as the conditions are very far removed from that, and the introduction of a dilution factor will not help the situation.

Consider the secular equilibrium between the pth and $(p+1)$th stage of ionization of the atom, and let there be N ions per cm.^{3} in both stages together, of which the fraction x is in the $(p+1)$th stage and $(1-x)$ in the pth stage. Let there be N_e electrons per cm.^{3}, and let the number of transitions per cm.^{3} per sec. be as follows:—

$N(1-x)Q$	photoelectric ionizations,
αNxN_e	photoelectric recombinations,
$N(1-x)N_eS_{12}$	ionizations by inelastic collisions with electrons,
$NxN_e^2S_{21}$	collision recombinations (three-body).

---

* B. Edlén, *Ark. Mat. Astr. Fys.*, **28**, No. 1B, 1942.
† H. Alfv n, *Ark. Mat. Astr. Fys.*, **27**, No. 25A, 1941.
‡ D. F. Martyn, *Nature*, **158**, 632, 1946.
§ K. O. Kiepenheuer, *M.N.*, **106**, 515, 1946.

The quantity Q depends on the radiation field, while the recombination coefficient α and the collision coefficients S_{12} and S_{21} depend on the agitation temperature.

In full equilibrium both

$$N(1-x)Q=\alpha NxN_e \tag{1.1}$$

and

$$N(1-x)N_eS_{12}=NxN_e^2S_{21}, \tag{1.2}$$

and in secular, but not necessarily full, equilibrium

$$N(1-x)(Q+N_eS_{12})=NxN_e(\alpha+N_eS_{21}) \tag{1.3}$$

or

$$\frac{N_ex}{1-x}=\frac{Q+N_eS_{12}}{\alpha+N_eS_{21}}. \tag{1.4}$$

In full equilibrium we have Saha's relation

$$\left(\frac{N_ex}{1-x}\right)_T=2(2\pi mkT)^{\frac{3}{2}}h^{-3}\frac{g_0(p+1)}{g_0(p)}e^{-\chi/kT},$$

where the g's are statistical weights, and χ the ionization potential. Equations (1.1) and (1.2) can be written

$$\frac{S_{12}}{S_{21}}=\frac{Q_T}{\alpha}=\left(\frac{N_ex}{1-x}\right)_T. \tag{1.5}$$

The observed high ionization in the corona may be caused by an excess of quanta (over normal solar radiation) in the soft X-ray spectrum, or by ionization by collision with fast electrons and atoms (or by both processes). We will inquire whether ionization by collision is a sufficient cause, and to do this we set Q (which depends on the radiation field) equal zero in (1.4), giving

$$\frac{N_ex}{1-x}=\frac{N_eS_{12}}{\alpha+N_eS_{21}}. \tag{1.6}$$

Ionization by collision has been discussed by J. J. Thomson * and R. H. Fowler †, and, in a few specific cases, by later writers using the quantum mechanics. References to these will be found in Mott and Massey's *Theory of Atomic Collisions*. According to Fowler the cross-section for ionization by electrons with energy ϵ is

$$\frac{\pi be^4}{\epsilon}\left(\frac{1}{\chi}-\frac{1}{\epsilon}\right),$$

where χ is the ionization potential, b a numerical factor of order unity, and e the electronic charge. This may be written

$$4b\pi a_0^2\left(\frac{\chi_H}{\chi}\right)^2\frac{\chi}{\epsilon}\left(1-\frac{\chi}{\epsilon}\right), \tag{1.7}$$

where a_0 is the radius of the first Bohr orbit, and the ionization potential of hydrogen is $\chi_H=e^2/2a_0$. With $b=\frac{1}{2}$ this agrees with the experimental curves for *H* and *He* shown in Mott and Massey's Figs. 31 and 32 in the range where ϵ is only slightly greater than χ. In the corona the mean value of ϵ/χ is less than one, and therefore most of the electrons capable of causing an ionization will have only a small energy excess. The corresponding value for S_{12} from Fowler's equation (38) is

$$S_{12}=\frac{4b\pi e^4}{\sqrt{2\pi}\chi^2}\left(\frac{kT}{m}\right)^{\frac{1}{2}}\exp(-\chi/kT), \tag{1.8}$$

* J. J. Thomson, *Phil. Mag.*, **23**, 449, 1912.
† R. H. Fowler, *Phil. Mag.*, **47**, 257, 1924.

which for 10^6 degrees becomes

$$S_{12} = 2{\cdot}0 \times 10^{-5} \frac{1}{\chi^2} \exp(-\chi/86) \qquad \text{(with } \chi \text{ in volts).}$$

The value of α for *Fe* XIV and other high-energy ions occurring in the corona is not known. Values for *O* calculated by Bates, Buckingham, Massey and Unwin * are given as a function of temperature in their Table II and those values can be extrapolated logarithmically to 10^6 degrees. If we assume α to be proportional to the square of Z (the ionic charge) as in Kramers' formula † we obtain $\alpha = 1{\cdot}5 \times 10^{-12}$ for *Fe* XIV at 10^6 degrees. This value of α represents a recombination collision cross-section of $2{\cdot}4 \times 10^{-21}$ cm.^{2}. With $N_e = 10^3$, using (1.5), one can show $\alpha \gg N_e S_{21}$. Then from (1.6)

$$\frac{x}{1-x} = \frac{S_{12}}{\alpha}. \tag{1.9}$$

Thus *the ionization is independent of electron pressure.* This has a simple physical explanation: Ionization by electron impact and radiative recombinations are both proportional to electron pressure, so that in a balance between the two the electron pressure drops out.

Calculations from (1.8) and (1.9) with $b = \frac{1}{2}$, $\alpha = 1{\cdot}5 \times 10^{-12}$, $T = 10^6$ degrees are shown in Table I.

TABLE I

Ionization by electron impact in the Corona

χ/kT	χ (volts)	$\frac{x}{1-x}$
3	258	10·0
3·5	301	4·4
4	344	2·1
4·5	387	1·00
5	430	0·48
5·5	473	0·24
6	516	0·12

Edlén's spectroscopic evidence ‡ indicates a maximum abundance of ions with an excitation potential of about 400 volts; that is to say $x = 1 - x$, where $\chi = 400$ volts. Table I shows that this agrees with the constant chosen. Hence molecular agitation corresponding to 10^6 degrees will account for the high ionization of the corona by electronic impact. Neglect of Q in (1.4) is therefore justified, and there is no evidence of soft X-ray emission from the photosphere in quantities large enough to influence the degree of ionization. There is also no physical reason for expecting the photosphere or chromosphere of themselves to emit soft X-rays and hence we might safely conclude that the *high ionization is caused by electronic impact.*

The degree of ionization in the corona has been discussed by Alfvén §, who also considers that ionization is mainly by collision. However, instead of introducing radiative recombinations to make up the equilibrium, an appeal is made to the finite life of the coronal structures. He finds that matter in the corona would be ionized up to 900 volts, which is high in comparison with the observed maximum at 400 volts.

* Bates, Buckingham, Massey and Unwin, *Proc. Roy. Soc.* A, **170**, 322, 1939.
† H. A. Kramers, *Phil. Mag.*, **46**, 836, 1923.
‡ *Loc. cit.*
§ *Loc. cit.*

2. *Line Intensities and Chemical Composition.*—Measurements of coronal line intensities may be used to determine the number of atoms in each ion, and by allowance for the ions that cannot be detected one may further determine the relative abundance of atoms. The number of hydrogen atoms may be inferred from the electron concentration, and hence it is possible to evaluate the hydrogen to metal ratio without having to adopt a critical excitation temperature to convert observed high-level hydrogen populations into numbers in the ground state. Since our purpose will be to compare the intensities of coronal lines with the coronal continuum, it is convenient to express the line intensities as equivalent breadths of the continuous spectrum scattered by electrons.

Table II gives the adopted equivalent breadths $\Delta\lambda$ of the emission lines in the lower corona. They are averaged from measurements by Lyot* and Grotrian†, and Allen‡, and from intensity estimates summarized by Mitchell.§ Other data in this table are taken from Edlén's original paper‖ and the reviews of his later work by Swings.¶

The light scattered by electrons in all directions per cm.^{3} per sec. within the wave-length range $\Delta\lambda$ is

$$\sigma 4\pi D F_\lambda \Delta\lambda N_e, \tag{2.1}$$

where σ is the electron scattering coefficient $\frac{8\pi}{3}\left(\frac{e^2}{mc^2}\right)^2 = 0{\cdot}66 \times 10^{-24}$, D is the dilution factor, F_λ the mean intensity of solar radiation per steradian per wave-length range, and N_e the number of electrons per cm.^{3}.

The light emitted in a spectrum line is

$$N_u A hc/\lambda, \tag{2.2}$$

where there are N_u atoms per cm.^{3} in the upper level, A is the transition probability per sec., and hc/λ the energy of the quantum.

If an excitation temperature T be assumed then

$$N_u = N_0 \frac{g_u}{g_0} e^{-E/kT}, \tag{2.3}$$

where there are N_0 atoms per cm.^{3} in the ground level, g_u and g_0 are the statistical weights, and E the upper excitation potential. If we now make $\Delta\lambda$ the equivalent breadth of an emission line then (2.1), (2.2) and (2.3) give

$$\frac{N_0}{N_e} = \frac{4\pi\sigma}{hc} D L \exp(E/kT), \tag{2.4}$$

where

$$L = \frac{g_0}{g_u} \cdot \frac{\lambda \Delta\lambda F_\lambda}{A}$$

Values used for this equation are given in Table II. In the case of *Fe* XIII, *Ca* XV and *Ni* XV the effective weight of the ground level has been doubled to allow for the population in near ground levels.

In order to determine the excitation temperature it is convenient to plot $\log_{10} L$ against the upper excitation potential in volts. This is done in Fig. 1. For the four ions which emit more than one spectrum line the slopes average about 0·4, and from a well-known relation the temperature becomes 5040/0·4,

* B. Lyot, *M.N.*, **99**, 580, 1939.
† W. Grotrian, *Z. Astrophys.*, **7**, 26, 1933.
‡ C. W. Allen, *M.N.*, **106**, 137, 1946.
§ S. A. Mitchell, *Handb. Astrophys.*, **7**, 401, 1934.
‖ *Loc. cit.*
¶ P. Swings, *Ap. J.*, **98**, 116, 1943 (Edlén's paper *Z. Astrophys.*, **22**, 30, 1942, is not available); *Pub. A.S.P.*, **57**, 117, 1945.

TABLE II

Line Intensities and Abundance of Ions in the Corona

λ	Element	Multiplet	E.P. (Upper) Volts	I.P. (Lower) Volts	A sec.$^{-1}$	$\Delta\lambda$ A.	F_λ*	$\frac{\lambda\Delta\lambda F_\lambda}{A}$ erg cm.$^{-1}$ sterad^{-1}	$\frac{g_0}{g_u}$	$\log\frac{N_0}{N_e}$
3328	*Ca* XII	$^2P_{\frac{3}{2}}-{}^2P_{\frac{1}{2}}$	3·72	589	488	0·6	1·2	0·05	$\frac{4}{2}$	$\bar{8}\cdot64$
3388	*Fe* XIII	$^3P_2-{}^1D_2$	5·96	325	87	11	1·3	6·1	$\frac{1(+1)}{5}$	$\bar{6}\cdot91$
3601	*Ni* XVI	$^2P_{\frac{1}{2}}-{}^2P_{\frac{3}{2}}$	3·44	455	193	1·2	1·5	0·34	$\frac{2}{4}$	$\bar{8}\cdot76$
3643	*Ni* XIII	$^3P_1-{}^1D_2$	5·82	350	18	0·1	1·5	0·30	$\frac{5}{5}$	$\bar{7}\cdot95$
3987	*Fe* XI	$^3P_1-{}^1D_2$	4·68	261	9·5	0·2	2·0	1·7	$\frac{5}{5}$	$\bar{6}\cdot25$
4086	*Ca* XIII	$^3P_2-{}^3P_1$	3·03	655	319	0·2	2·5	0·06	$\frac{5}{3}$	$\bar{8}\cdot36$
4231	*Ni* XII	$^2P_{\frac{3}{2}}-{}^2P_{\frac{1}{2}}$	2·93	318	237	0·6	2·8	0·26	$\frac{4}{2}$	$\bar{7}\cdot02$
4359	? *A* XIV	$^2P_{\frac{1}{2}}-{}^2P_{\frac{3}{2}}$	2·84	682	108	0·1	2·9	0·12	$\frac{2}{4}$	$\bar{8}\cdot05$
5116	*Ni* XIII	$^3P_2-{}^3P_1$	2·42	350	157	0·6	3·1	0·6	$\frac{5}{3}$	$\bar{7}\cdot12$
5303	*Fe* XIV	$^2P_{\frac{1}{2}}-{}^2P_{\frac{3}{2}}$	2·34	355	60	28	3·0	75	$\frac{2}{4}$	$\bar{6}\cdot67$
5536	*A* X	$^2P_{\frac{3}{2}}-{}^2P_{\frac{1}{2}}$	2·24	421	106	0·1	2·9	0·15	$\frac{4}{2}$	$\bar{8}\cdot53$
5694	*Ca* XV	$^3P_0-{}^3P_1$	2·18	814	95	0·2	2·8	0·34	$\frac{1(+1)}{3}$	$\bar{8}\cdot40$
6374	*Fe* X	$^2P_{\frac{3}{2}}-{}^2P_{\frac{1}{2}}$	1·94	233	69	4	2·6	9·5	$\frac{4}{2}$	$\bar{6}\cdot21$
6702	*Ni* XV	$^3P_0-{}^3P_1$	1·85	422	57	1·5	2·3	4·0	$\frac{1(+1)}{3}$	$\bar{7}\cdot34$
7060	*Fe* XV	$^3P_1-{}^3P_2$	31·7	390		1·0	2·1	15	$\frac{1}{5}$	—
7892	*Fe* XI	$^3P_2-{}^3P_1$	1·57	261	44	7	1·7	21	$\frac{5}{3}$	$\bar{6}\cdot34$
8024	*Ni* XV	$^3P_1-{}^3P_2$	3·39	422	22	0·3	1·6	1·7	$\frac{1(+1)}{5}$	$\bar{7}\cdot34$
10747	*Fe* XIII	$^3P_0-{}^3P_1$	1·15	325	14	60	0·9	410	$\frac{1(+1)}{3}$	$\bar{5}\cdot07$
10798	*Fe* XIII	$^3P_1-{}^3P_2$	2·30	325	9·7	40	0·9	390	$\frac{1(+1)}{5}$	$\bar{5}\cdot27$
5303	Outer Corona					2·2	3·0	5·8	$\frac{2}{4}$	$\bar{5}\cdot09$

* In units of 10^{14} erg cm.$^{-3}$ sec.$^{-1}$ sterad^{-1}.

i.e. 13,000 deg. K. This compares with the excitation temperatures 29,000 deg. K. and 21,000 deg. K. obtained by Kun Huang.* It is evident that this rather low excitation temperature could not be applied to the 31·7 volt line 7060 A. of *Fe* XV.

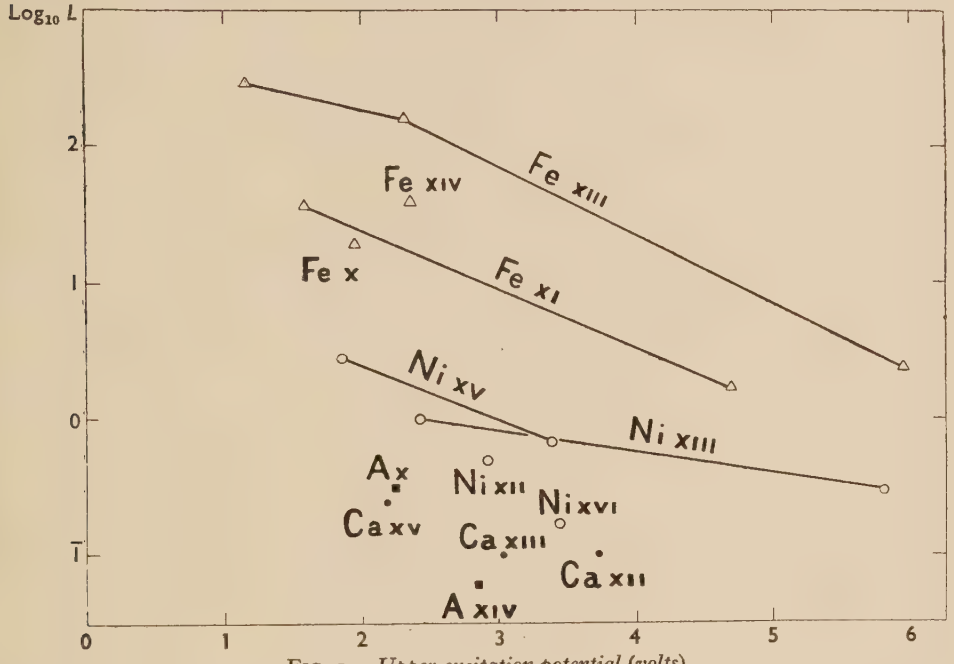


FIG. 1.—*Upper excitation potential (volts).*

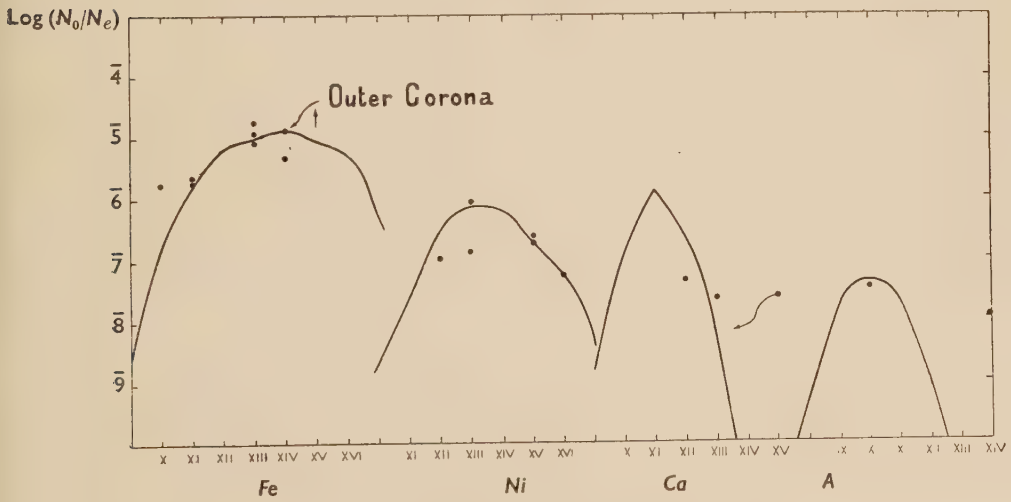


FIG. 2.—*Abundance of Fe, Ni, Ca and A ions.*
· *Observed values.*
— *Calculated from Table I.*

Using $4\pi\sigma/hc = 4{\cdot}2 \times 10^{-8}$, and $D = \frac{1}{3}$ for the inner corona, we obtain from (2.4) the various values of

$$\log_{10}(N_0/N_e) = 0{\cdot}4\,E\,(\text{volts}) + \log_{10} L - 7{\cdot}85,$$

given in the last column of Table II. These values of the abundance of the various ions in relation to electrons are plotted in Fig. 2. The relative populations

* Kun Huang, *Ap. J.*, **101**, 187, 1945.

of adjacent ions from Table I have been weighted in accordance with the ground statistical weight (g_0 in Table II) and are shown as the calculated relative ionizations in Fig. 2. The level of the calculated curve for *Fe*, *Ni*, *Ca* and *A* has been adjusted to fit the observed points. No change in the adopted values of S_{12}/α combined with the electron temperature of 10^6 degrees would improve the fit appreciably, and hence we have obtained some confirmation for these values. The variations of temperature that undoubtedly occur throughout the corona * can account for the occasional appearance of lines (e.g. 5694 A. of *Ca* XV) which according to the calculations should not be visible.

In order to determine the total number of *Fe*, *Ni*, *Ca* and *A* atoms, it is necessary to estimate the abundance of all ions, including those that emit no observable spectrum. We use the theoretical values for this purpose after having fitted them to the experimental values as well as possible. Adding for all the ions, we obtain the numbers of atoms per electron in the corona given in Table III. The inclusion of $H=0{\cdot}75$ is obtained by assuming an $H:He$ number ratio of 6:1.

TABLE III

Atoms per electron in the Corona

H	*Fe*	*Ni*	*Ca*	*A*
0·75	$4{\cdot}7\times10^{-5}$	$1{\cdot}9\times10^{-6}$	$1{\cdot}6\times10^{-6}$	8×10^{-8}

When both *H* and *He* are fully ionized, there will be eight electrons to six hydrogen nuclei, since the abundance of other elements is negligible.

For comparison, Table IV contains the relative cosmic abundance of elements 1–30 as obtained from various other sources.† All recent estimates are in agreement with the coronal observations in giving a high hydrogen to metal ratio. It will be remembered that the chief reason for the lower ratio used in Unsöld's *Physik der Sternatmosphären* was Strömgren's evidence from the internal constitution of stars. However, the admission of a finite helium content in stars has the effect ‡ of increasing the hydrogen ratio, and the more recent measurements by Menzel §, Unsöld ||, and Aller ¶, indicate sufficient helium. Hoyle's analysis ** shows that a very small proportion of metals to hydrogen can be brought into agreement with observations. The various values have been averaged in the last column of Table IV to give what might be regarded as a present-day estimate of the composition (log relative number of atoms) of stellar atmospheres. The ratio of groups of elements is given in Table V.

Since the various estimates of composition set out in Table IV are not always independent, one cannot determine the probable errors of the average value. However, the general agreement for different estimates suggests the probable error of the mean log abundance would not be greater than 0·2 for most elements. In the corona the probable error would be about 0·3 for *H*, *Fe* and *Ni* and greater for *Ca* and *A*. The abundance figures for the corona, in Table IV, are fitted

* M. Waldmeier, *Ast. Mit. Eidg. St.*, Zürich, Nr. 149, 1947.

† C. H. Payne, *Stellar Atmospheres*, p. 184, 1925. H. N. Russell, *Ap. J.*, **70,** 11, 1929; **75,** 337, 1932; *cf.* also L. Goldberg and L. H. Aller, *Atoms, Stars and Nebulae*, p. 114, 1943. L. H. Aller and D. H. Menzel, *Ap. J.*, **102,** 239, 1945. J. L. Greenstein, *Ap. J.*, **95,** 161, 1942. F. H. Loring, *The Chemical Elements*, p. 32, 1923. F. G. Watson, *Between the Planets*, p. 177, 1941.

‡ B. Strömgren, *Ap. J.*, **87,** 520, 1938.

§ D. H. Menzel, *Pub. Lick Obs.*, **17,** 281, 1931.

|| A. Unsöld, *Physik der Sternatmosphären*, p. 82, 1938; *Ap. J.*, **100,** 105, 1944; *Akad. Wiss. Gott.*, p. 70, 1945.

¶ L. H. Aller, *Pub. A.S.P.*, **59,** 144, 1947; *Ap J.*, **104,** 347, 1946.

** F. Hoyle, *M.N.*, **106,** 255, 1947.

TABLE IV

Cosmic Abundance (Log Relative Number of Atoms)

	Payne	Russell		Menzel		Unsöld		Aller	Aller and Menzel	Green-stein	Earth's Crust	Meteors	Corona	Mean
1 *H*	12·9	11·5	9·9	13·6	12·1	9·7	11·6	12·0	11·9	11·9	8·3		11·8	11·8
2 *He*	10·2		8·4		11·4		10·9	10·7	10·9					11·0
3 *Li*	1·9	2·0									5·2			4
4 *Be*		1·8									4·3	3		3
5 *B*		4·5									4·3			4·4
6 *C*	6·4	7·4			7·7		7·9	7·8	8·7	7·8	6·3	6·0		7·6
7 *N*		7·6			8·2		8·2	8·0	8·2	8·3	5·7			8·0
8 *O*	8·0	9·0	8·4		8·7	8·3	8·6	8·6	8·3		8·5	8·3		8·5
9 *F*		6·0							4·9		5·5	4·0		5·1
10 *Ne*							8·6	7·8	6·9					7·8
11 *Na*	7·1	7·2	7·2	7·2	6·7	6·6					7·3	6·0		7·0
12 *Mg*	7·8	7·8	7·8	7·4	8·5	7·5	7·3	7·6			7·2	7·6		7·6
13 *Al*	6·9	6·4	6·4	6·2	6·5	7·1	6·1	6·3			7·7	6·3		6·6
14 *Si*	7·2	7·3	7·3	5·7	8·0	7·8	7·3	7·5		7·5	8·2	7·7		7·4
15 *P*		4·0						5·9			5·9	5·5		5·8
16 *S*		5·7			7·7		6·8	6·8	7·4	6·7	5·8	6·8		6·7
17 *Cl*									6·2		6·1	4·7		5·7
18 *A*									6·1				4·8	5·5
19 *K*	5·4	6·8	6·8		5·2						7·1	5·2		6·1
20 *Ca*	7·0	6·7	6·7	6·6	6·7	6·9	6·1			6·1	7·2	6·4	6·1	6·6
21 *Sc*		3·6		3·8							2	3		3·4
22 *Ti*	6·0	5·2		5·7	4·7						6·4	5·2		5·5
23 *V*	4·9	5·0		5·6	4·2						5·2	4·0		5·0
24 *Cr*	5·9	5·7	6·1	6·1	5·0						5·4	5·4		5·6
25 *Mn*	6·5	5·9		6·4	5·2						5·6	5·5		5·8
26 *Fe*	6·7	7·2	7·2	7·0	7·1	7·9	7·6				7·2	7·6	7·6	7·4
27 *Co*		5·6		5·0	4·8						4·0	5·5		5·0
28 *Ni*		6·0		6·0	6·5	6·7					5·0	6·6	6·2	6·1
29 *Cu*		5·0			4·5						4·5	4·4		4·6
30 *Zn*	6·1	4·9		4·3	5·7						4·1			5·0

to the mean of the other columns for *Fe*, *Ni* and *Ca*, and on this scale the abundance of *H* is the same as the mean of other astrophysical determinations.

Spectroscopic determinations of the composition of stellar atmospheres are normally based on the rule that the sum of the oscillator strengths from a particular level is equal to either unity or the number of valence electrons.

TABLE V

Abundance ratio in stellar atmospheres

	Hydrogen	Helium	Oxygen Group (*Li–Ne*)	Metals (*Na–Zn*)
By numbers	6000	1000	4	1
By weight	6000	4000	60	40

However, an evaluation of King's measurements * of absolute oscillator strengths (f-values) in *Fe* lead to Σf about 0·2. Thus the abundance of *Fe* in the Sun should be some five times as great as given in the Russell mixture.† Such a change would bring the iron abundance into better agreement with the corona. Unfortunately, however, it is not known whether a similar correction should be made for other elements. Some allowance for this correction has been made in determining the mean in Table IV.

3. *The Outer Corona.*—In the outer corona one can readily compute the degree of excitation without assuming an empirical excitation temperature. In this case the frequency of collision becomes much less than the mean life of the excited states and the excitation must be entirely due to solar radiation. Then the factor $e^{-E/kT}$ in (2.3) becomes $D \cdot e^{-E/kT_s}$, where T_s is the surface temperature of the Sun and D the dilution factor. Instead of (2.4) we now find D cancels and

$$N_0/N_e = (4\pi\sigma/hc)Le^{E/kT_s}. \tag{3.1}$$

Since we have shown in Section 1 that the degree of ionization is independent of pressure, this equation shows that above a certain height in the corona the equivalent breadth of emission lines should not change with height.

The intensity measurements of the coronal line 5303 A. made at the 1940 eclipse ‡ were expressed in equivalent angstroms of the continuous spectrum. The fraction of the continuous intensity due to electron scatter is $J_K/(J_K+J_F)$, which was also measured at the same eclipse (the component J_F is not true coronal light).

The line intensity measurements have been multiplied by $(J_K+J_F)/J_K$ in Table VI to give the equivalent breadth in terms of the electron scatter continuum as a function of ρ, the distance from the Sun's centre in solar radii.

TABLE VI

Equivalent breadth of 5303 in terms of the electron scatter continuum

ρ	1·2	1·4	1·6	1·8	2·0	2·2	2·4	2·6
Equivalent breadth (A.)	(10)	6·5	4·3	3·1	2·5	2·2	2·2	2·3

It is seen that beyond $\rho = 2{\cdot}0$ the equivalent breadth remains constant at 2·2 A. § Putting $T_s = 5900$ deg. K., (3.1) gives $(N_0/N_e) = 12 \times 10^{-6}$. This is in reasonable agreement with the value 5×10^{-6} for the inner corona, for which use was made of an empirical excitation temperature. The outer corona value is plotted in Fig. 2, and is included in the estimate of relative abundance discussed in Section 2.

4. *The Collision Cross-Section.*—We have attributed the variation of equivalent breadths of coronal emission lines with distance from the Sun to variation in the ratio of collision excitation to excitations by radiation. We can actually

* R. B. King, *Ap. J.*, **95**, 78, 1942.

† H. N. Russell, *Ap. J.*, **70**, 11, 1929.

‡ C. W. Allen, *M.N.*, **106**, 137, 1946.

§ The constancy of the intensity of 5303 A. relative to the light scattered by electrons from $\rho = 2{\cdot}0$ to $\rho = 2{\cdot}6$ gives strong support to the supposition that ionization is independent of electron pressure as required by equation (1.9), and therefore supports the theory that ionization is determined by a balance between collisions and recombination. If this is granted then the observed level of ionization confirms the ratio adopted for S_{12}/α. Hence, although values given in the text were not calculated from wave mechanics, there is good reason to believe that they cannot be greatly in error.

determine the collision excitation cross-section of *Fe* XIV 5303 A. by comparing it with the probability of radiation excitation. For this purpose we note from Table VI that at the height $\rho = 1{\cdot}6$ the coronal line intensity (i.e. equivalent breadth compared with light scattered by electrons) is about twice the value in the outer corona, and hence at this height the number of excitations by electron impact is equal to the number of excitations by solar radiation.

At $\rho = 1{\cdot}6$ the number of excitations by solar radiation per cm.^{3} per sec. is

$$N_{Fe\,XIV} \cdot \frac{g_u}{g_0} \cdot D \cdot 10^{-\frac{5040\,E}{T_s}} \cdot A, \tag{4.1}$$

where $N_{Fe\,XIV}$ is the number of *Fe* XIV atoms (in the ground state) per cm.^{3}. With $g_u/g_0 = 2$, $D = \frac{1}{2}\{1-(1-\rho^{-2})^{\frac{1}{2}}\} = 0{\cdot}11$, $E = 2{\cdot}34$ volts, $T_s = 5900$ deg. and $A = 60$ sec.$^{-1}$, we find 0·13 excitations of each *Fe* XIV atom per sec. by radiation.

The number of excitations by electron impact on the other hand is

$$N_{Fe\,XIV} N_e S(8kT/\pi m)^{\frac{1}{2}}, \tag{4.2}$$

per cm.^{3} per sec., where S is the effective cross-section. Inserting * $N_e = 9{\cdot}3 \times 10^6$ and $T = 10^6$ and equating (4.1) to (4.2), we get

$$S = 2{\cdot}3 \times 10^{-17}\ \text{cm.}^2 \text{ or } 0{\cdot}26\pi a_0^2,$$

where a_0 is the radius of the first hydrogen orbit. It will be noticed that the abundance of *Fe* is not involved in the result.

5. *Line Emission in the Ultra-violet.*—A knowledge of collision cross-section may be used to determine the amount of line emission in the ultra-violet, since one collision is necessary for every quantum emitted. Permitted transitions of ultimate spectrum lines are likely to be the strongest and it would not be safe to use the cross-section just determined which is for a forbidden magnetic dipole transition. We therefore adopt the formula given by Mott and Massey †,

$$Q_{on} \simeq \frac{64\pi^5 m^2 \epsilon^4}{k\,h^4} |\, x_{on}|^2 \log \frac{2mv^2}{E_h - E_0}$$

(in their own notation). We can introduce the transition probability A in place of the matrix elements $|\, x_{on}|^2$, and obtain the cross-section as a function of v the electron velocity

$$S_v = \pi e^4 A \frac{3mc^3}{8\pi^2 e^2 \nu^2} \cdot \frac{1}{h\nu} \cdot \frac{2}{mv^2} \log \frac{2mv^2}{h\nu}, \tag{5.1}$$

or, in terms of the oscillator strength f (and omitting the slowly varying term)

$$S_v = 3\pi e^4 f \frac{1}{h\nu} \cdot \frac{2}{mv^2}. \tag{5.2}$$

This formula agrees quite well with measurements for hydrogen with 1000 volt electrons (Mott and Massey, p. 183). From it we may calculate the total number of excitations in all (permitted) lines, since Σf (for the ground state) is equal to the number of valence electrons. However, the formula is only valid for fast electrons. Where the kinetic energy of the electron is comparable with the excitation energy, experimental work indicates that Fowler's formula ‡

$$S_v \propto \frac{2}{mv^2}\left(\frac{1}{h\nu} - \frac{2}{mv^2}\right),$$

---

* C. W. Allen, *M.N.*, **107**, 426, 1947.

† *Atomic Collisions*, p. 177.

‡ R. H. Fowler, *Phil. Mag.*, **47**, 257, 1924.

gives the right form of variation of S_v with v. We shall therefore assume

$$S_v = 3\pi e^4 f \frac{2}{mv^2}\left(\frac{1}{h\nu} - \frac{2}{mv^2}\right) \tag{5.3}$$

as a general formula for the collision excitations of permitted lines. Then if there are N atoms of a certain kind and N_e electrons per cm.^{3}, the number of excitations per cm.^{3}, by collision is

$$NN_e \frac{12\pi e^4 f}{(2\pi mkT)^{\frac{1}{2}}} \cdot \frac{1}{\chi'} \{\exp(-y) - yEi(y)\}, \tag{5.4}$$

where $y = \chi'/kT$ and χ' is the energy of the line. The value of N_e as a function of ρ can be written *

$$N_e(\rho) = 10^8(1{\cdot}55\rho^{-6} + 2{\cdot}99\rho^{-16}).$$

We may assume the mean dilution to be $\frac{1}{3}$, so that $\frac{2}{3}$ of the quanta emitted escape from the Sun. Then, supposing the emission from a particular ion to be concentrated in a single (ultimate) line for which $f = 1$ and the excitation energy is χ' the total number of permitted quanta emitted by this ion in the entire corona above ρ_0 is

$$Q = \frac{12\pi e^4}{(2\pi mkT)^{\frac{1}{2}}} \cdot \frac{1}{\chi'} \{\exp(-y) - yEi(y)\} p' M, \tag{5.5}$$

where $M = \frac{2}{3}4\pi r_\odot^3 \{0{\cdot}266\rho_0^{-9} + 0{\cdot}490\rho_0^{-19} + 0{\cdot}307\rho_0^{-29}\} \times 10^{16}$
$= 2{\cdot}3 \times 10^{49}$.

Here $r_\odot$ is the Sun's radius $6{\cdot}96 \times 10^{10}$ cm., and we use $\rho_0 = 1{\cdot}015$ for the base of the corona. We put $p = N/N_e$ and let p' refer to a particular ion. In the case of those spectra where a particular ion is designated, $p^2 = p(1-x)$, but in the general case of metals where both stages of ionization count alike, $p' = p$. With $T = 10^6$ and χ' in ergs this gives

$$Q = (p'/\chi')\{\exp(-y) - yEi(y)\} \times 5{\cdot}2 \times 10^{31} \text{ quanta per second.} \tag{5.6}$$

Values of $p'/\chi'\{\exp(-y) - Ei(y)\}$ for various ions are shown in Table VII.

TABLE VII

Contributions to Line Emission ($\chi' > 10$ volts)

Ion	p	$\frac{x}{1-x}$	p'	χ' volts	y	$(p'/\chi')\{\exp(-y) - yEi(y)\}$ χ' in ergs	Q 10^{35} sec.$^{-1}$	$E = Q\chi'$ 10^{24} erg sec.$^{-1}$
H I	0·75	7×10^7	$1{\cdot}1 \times 10^{-8}$	10·2	0·12	500	0·3	0·5
He II	0·12	7×10^5	$1{\cdot}7 \times 10^{-7}$	40·6	0·47	900	0·5	3·2
O VI	5×10^{-4}	10^3	5×10^{-7}	12	0·14	17000	9	17
Mg X	5×10^{-5}	2·8	$1{\cdot}2 \times 10^{-5}$	20	0·23	200000	110	350
Metals	$1{\cdot}2 \times 10^{-4}$	1·0	$1{\cdot}2 \times 10^{-4}$	200	2·32	10000	5	160
Total							125	530

In calculating the ionizations $x/(1-x)$ from (1.9) the recombination coefficient α has been taken as $1{\cdot}5 \times 10^{-12}$ for metals (from Section 1) and $1{\cdot}5 \times 10^{-12} \times Z^2/169$ for *O* VI and *Mg* X, while the coefficient for *H* has been taken † as $5{\cdot}4 \times 10^{-15}$, and 22×10^{-15} taken for *He*. The constant b has been taken as 2 for *H* and *He*, 1 for *O* and $\frac{1}{2}$ for metals. The excitation energy of the principal lines of *Mg* X

* C. W. Allen, *M.N.*, **107**, 426, 1947.

† G. Cillié, *M.N.*, **92**, 320, 1932.

$(2\,^2S_{\frac{1}{2}} - 2\,^2P_{\frac{1}{2},\frac{3}{2}})$ has been estimated by extrapolating terms of the lithium-like spectra*, and using the data† for *Na* IX. The average excitation potential for the metals generally has been taken equal to half the ionization potential of the best represented ions, i.e. those with ionization potentials near 400 volts. In ions with complete or nearly complete shells the excitation potential of the principal lines is about 0·8 of the ionization potential and only a small percentage of electrons have sufficient energy to cause an excitation. Ions with very incomplete shells make a much larger contribution to emission, since the excitation energy is smaller and electron collisions more effective. *Mg* X should make an outstanding contribution, since the principal lines have very low energy (only 20 volts) and it is a well-represented ion of an abundant element, its ionization potential being about 370 volts. The calculations, which of course are very rough, suggest that the greater part of coronal line emission comes from this one ion whose principal lines lie near 610 A. The oscillator strength for the doublet has been assumed unity, but this may be doubted since the line energy is only 5 per cent of the ionization energy. Data concerning the *Mg* X spectrum, oscillator strengths, cross-sections and recombination coefficients are very desirable as they would decide whether this ion dominates the coronal line emission spectrum as suggested.

The permitted transitions just discussed will have excitations greater than 10 volts. It would not be expected that forbidden transitions beyond 10 volts (namely transitions from one configuration to another) would be as frequent as the permitted transitions. Hence the contributions to line emission in Table VII would not have to be increased appreciably to allow for forbidden transitions. There are, however, many excitations caused by forbidden transitions within configurations. Most of these are in the visible spectrum—the coronal lines—and only a negligible quantity (e.g. from *Si* IX) would reach 10 volts. If we restrict our interest to the far ultra-violet beyond 10 volts (1240 A.) the estimates given in Table VII should be complete.

6. *Continuous Emission in the Ultra-violet.*—The number of recombinations each emitting an ultra-violet quantum in a continuous spectrum may be computed in a similar manner. If p is the number of ions per electron, then, as in (1.1), the number of quanta emitted per cm.^{3} per sec. will be $pN_e^2\alpha$. The number of quanta emitted outwards from the whole corona per sec. becomes

$$Mp\alpha = 2{\cdot}3 \times 10^{49} p\alpha. \qquad (6.1)$$

Adopting the chemical composition of Section 2 and the α of Section 5 we obtain Table VIII.

TABLE VIII

Element	Z	p	pZ^2	α cm.^{3} sec.$^{-1}$	$p\alpha$	χ_m volt	$p\alpha\,\chi_m$ cm.^{3} sec.$^{-1}$ volt
H	1	0·75	0·75	$5{\cdot}4 \times 10^{-15}$	$4{\cdot}0 \times 10^{-15}$	66	$2{\cdot}7 \times 10^{-13}$
He	2	0·12	0·48	$22{\cdot}0 \times 10^{-15}$	$2{\cdot}6 \times 10^{-15}$	124	$3{\cdot}2 \times 10^{-13}$
O Group	6	0·0005	0·02	$0{\cdot}4 \times 10^{-12}$	$0{\cdot}2 \times 10^{-15}$	218	$0{\cdot}4 \times 10^{-13}$
Metals	13	0·00012	0·02	$1{\cdot}5 \times 10^{-12}$	$0{\cdot}2 \times 10^{-15}$	486	$1{\cdot}0 \times 10^{-13}$
Total		0·87	1·27		$7{\cdot}0 \times 10^{-15}$		$7{\cdot}3 \times 10^{-13}$

* B. Edlén and A. Ericson, *Zs. f. Phys.*, **64**, 64, 1930.
† J. Söderquist, *Reg. Soc. Sci. Upsal.*, **9**, No. 7, 1934.

From (6.1) and Table VIII the number of quanta emitted from the whole corona is

$$1{\cdot}6 \times 10^{35} \text{ quanta sec.}^{-1}.$$

The energy of each quantum is somewhat greater than the ionization potential and the mean energy χ_m based on the λ^3 law is given in Table VIII. The free-bound energy emitted by the whole corona is

$$Mp\alpha\chi\,(\text{ergs}) = 2{\cdot}7 \times 10^{25} \text{ erg sec.}^{-1}.$$

The free-free emission may be calculated as follows. If there are N_e electrons and N ions of a certain kind per cm.^{3}, the ionic charge being Z, the total energy emitted in the free-free spectrum * is

$$NN_eBZ^2T^{\frac{1}{2}} \text{ erg cm.}^{-3} \text{ sec.}^{-1}, \tag{6.2}$$

where

$$B = \frac{128\pi^3 e^6 k^{\frac{1}{2}}}{(6\pi m)^{\frac{3}{2}} hc^3} = 1{\cdot}42 \times 10^{-27}.$$

The emission from the whole corona is then

$$\Sigma pZ^2 \,.\, B \,.\, M \,.\, T^{\frac{1}{2}} = 4{\cdot}1 \times 10^{25} \text{ erg sec.}^{-1}.$$

The emission is proportional to $\exp(-h\nu/kT)$ and thus the energy above 10 volts is

$$3{\cdot}6 \times 10^{25} \text{ erg sec.}^{-1}.$$

The mean energy of quanta beyond 10 volts is $0{\cdot}53\,kT$ or $7{\cdot}3 \times 10^{11}$ erg. Thus the free-free quantum emission is

$$5{\cdot}0 \times 10^{35} \text{ quanta sec.}^{-1}.$$

7. *Significance of the Coronal Ultra-violet Emission.*—The importance of the ultra-violet emission from the corona depends in many connections upon whether it greatly exceeds in intensity emission in the same frequencies from other regions in the Sun. On account of the great uncertainty which was felt about remote ultra-violet solar emission before coronal theory made its recent advances, it was usual to take refuge in a purely arbitrary representation of the remote solar spectrum and suppose it to be that of a black-body at 6000 deg. Advance in knowledge about the corona now enables us to go a step further. We do not propose to treat the matter in detail in the present paper, but merely to present a brief sketch of a scheme in relation to which the importance of the coronal emission may be assessed.

If we divide the solar atmosphere into three regions, photosphere, chromosphere and corona, then there is hardly any doubt that in the far ultra-violet the emission from the photosphere is equal to that of a black-body at a temperature much nearer 5000 deg. than 6000 deg. (In the very important hydrogen Lyman continuum the absorption coefficient of the solar atmosphere is much higher than it is in the visible spectrum, so that we do not "see down to" a great depth, in fact hardly at all below the surface at 4830 deg.) But the chromosphere is observed to be at a much higher temperature. Thus Redman † gives 30,000 deg. for the kinetic temperature, and Cillié and Menzel ‡ give 10,000 deg. for the electron temperature of the chromosphere. Now the optical thickness of the chromosphere is quite considerable in the Lyman continuum (of order unity),

* G. Cillié, *M.N.*, **92**, 820, 1932.
† R. O. Redman, *M.N.*, **102**, 140, 1942.
‡ G. G. Cillié and D. H. Menzel, *Harv. C. O. Circ.*, No. 410, 1935.

so that it is capable of radiating an amount equal to that of a black-body at a temperature much greater than 5000 deg. It is not easy to decide the levels to which the observed temperatures refer, and if a value between 5000 deg. and say 10,000 deg. were adopted for the effective temperature of the chromosphere in the Lyman continuum the temperature would enter very critically in the resultant emission computed from it. It seems at present safer to attack the problem in another way and say that the (excess) chromospheric emission cannot exceed in energy the amount of heat transferred from the corona inwards by conduction, as examined by Alfvén.*

A preliminary examination in this way has given a value of about 6500 deg. for the effective temperature of the chromosphere, yielding $2{\cdot}0 \times 10^{37}$ quanta $sec.^{-1}$ from the whole of the Sun in the Lyman continuum. The remote ultra-violet solar spectrum is therefore supposed to be that of a black-body at 5000 deg. for energies just inferior to 13·6 volts, the Lyman limit: to rise sharply at that limit giving the Sun a Lyman "appendage" (similar to but much stronger than the Balmer appendages observed in some B type stars): to decline thereafter like a black-body at 6500 deg., and finally to show some emission lines, the principal lines of *Mg* X being the strongest. The total number of quanta emitted from the corona is less than the number of Lyman quanta emitted by the chromosphere.

This scheme fits in very well with our knowledge of the ionosphere, and makes the identification of the E region with electrons liberated by O_2, and of the two F regions with electrons liberated by O, N and N_2 almost certain. Since the radiation ionizing O_2 (ionization potential 12·2 volts) has an effective temperature of only 5000 deg. the scheme explains why the E region receives less quanta for its maintenance than the F regions. This puzzling fact could not easily be explained by any smooth black-body curve without an appendage. The appendage corresponds in fact to the postulates of Saha † and his co-workers, that the Sun exhibits a bright line ultra-violet spectrum forming the F region. A bright continuum and certain bright lines now appear as a natural consequence of coronal agitation.

The authors hope to give a more detailed treatment of chromospheric emission and the connection between corona, chromosphere and ionosphere in a further paper.

The authors wish to acknowledge their indebtedness to Dr S. C. B. Gascoigne for valuable discussions on collision cross-sections.

Commonwealth Observatory,
Mount Stromlo,
Canberra :
1947 *November* 15.

* *Loc. cit.*
† M. N. Saha, *Proc. Roy. Soc.* A, **160**, 155, 1937.